Electronic structure, defect formation and passivation of 2D materials

Haichang Lu

St. Edmund's college

Department of Engineering

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

I would like to dedicate this thesis to my loving parents Zhongfei Lu and Lanfen Chen.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains less than 65,000 words including appendices, bibliography, footnotes, tables and equations and has less than 150 Figures.

Permission is granted to consult the information and results contained within this thesis for the purpose of private study only, not for publication.

This thesis contains roughly 37700 words and 66 Figures.

Haichang Lu

June 27th 2018

Acknowledgment

I want to thank many people for helping me go through my 4 years Ph.D. smoothly and with a satisfying end.

The first person and the person I want to thank most is my supervisor, Professor John Robertson. He chose me as one of his group members and led me to a very promising field, the 2D materials. The projects he assigned me were interesting, challenging and also rendering productivity. His way of guidance is inspiring and he affected me profoundly and substantially in the aspect of tracking the most important issue and solving it in my research. His physical instinct is outstanding and his sense of what is important in rapid changing scientific fashion is incisive. I am also grateful for his financial support for my conference trip to the USA. I also need to thank Professor Stewart Clark from Durham University for his support about the use of the DFT code CASTEP. He is always kind and patient in answering my questions.

I would also like to thank my colleagues in the department of engineering, particularly Dr. Yuzheng Guo, who was the senior research associate when I entered Cambridge. He provided me with lots of technical support for doing the DFT calculation. Besides, he inspired me and encouraged me to be a junior researcher. Also, I want to thank Dr. Hongfei Li, Dr. Xiaoming Yu, Prof. Huanglong Li and Dr. Zhigang Song as the senior colleagues in the group to give me advice on research. For experimentalists, I want to thank Mrs. Shan Zheng as a collaborator and Dr. Xingyi Wu and Mr. Guandong Bai for introducing experimental details. I appreciate the daily discussion from Miss Han Zhang and Dr. Zhigang Song.

During my Ph.D. I also received countless of help from my previous supervisors and colleagues, I would like to thank specially to Prof. Guangshan Tian for his advice on academic career and I would also like to thank Miss Yifan Ma and Miss Jinghang Yang for their encouragement.

Finally, I want to express my great gratitude to my parents for their financial support in my first year and for travel expense. This thesis is dedicated to them.

Abstract

Electronic structure, defect formation and passivation of 2D materials

Haichang Lu

The emerging 2D materials are potential solutions to the scaling of electronic devices to smaller sizes with lower energy cost and faster computing speed. Unlike traditional semiconductors e.g. Si, Ge, 2D materials do not have surface dangling bonds and the short-channel effect. A wide variety of band structure is available for different functions. The aim of the thesis is to calculate the electronic structures of several important 2D materials and study their application in particular devices, using density functional theory (DFT) which provides robust results.

The Schottky barrier height (SBH) is calculated for hexagonal nitrides. The SBH has a linear relationship with metal work function but the slope does not always equal because Fermi level pinning (FLP) arises. The chemical trend of FLP is investigated. Then we show that the pinning factor of Si can be tuned by inserting an oxide interlayer, which is important in the application to dopant-free Si solar cells.

Apart from contact resistance, we want to improve the conductivity of the electrode. This can be done by using a physisorbed contact layer like $FeCl_3$, $AuCl_3$, and SbF_5 etc. to dope the graphene without making the graphene pucker so these dopants do not degrade the graphene's carrier mobility.

Then we consider the defect formation of 2D HfS_2 and SnS_2 which are candidates in the n-type part of a tunnel FET. We found that these two materials have high mobility but there are also intrinsic defects including the S vacancy, S interstitial, and Hf/Sn interstitial.

Finally, we study how to make defect states chemically inactive, namely passivation. The S vacancy is the most important defect in mechanically exfoliated 2D MoS_2. We found that in the most successful superacid bis(trifluoromethane) sulfonamide (TFSI) treatment, H is the passivation agent. A symmetric adsorption geometry of 3H in the -1 charge state can remove all gap states and return the Fermi level to the midgap.

Publication List

[1] **H Lu,** Y Guo and J Robertson, Chemical trends of Schottky barrier behavior on monolayer hexagonal B, Al, and Ga nitrides. *J. Appl. Phys*, **120**, 065302 (2016).

[2] **H Lu**, Y Guo and J Robertson, Charge transfer doping of graphene without degrading carrier mobility. *J. Appl. Phys*, **121**, 224304 (2017).

[3] **H Lu**, Y Guo and J Robertson, Band edge States, Intrinsic Defects and Dopants in Monolayer HfS_2 and SnS_2, *Appl. Phys. Lett*, **112**, 062105 (2018).

[4] S Lee, A Nathan, J Alexander-Webber, P Braeuninger-Weimer, A Sagade, **H Lu**, D Hasko, J Robertson, S Hofmann, Dirac-Point Shift by Carrier Injection Barrier in Graphene Field-Effect Transistor Operation at Room Temperature. Accepted by *ACS Appl. Mater. Interfaces*, **10**, 10618 (2018).

[5] **H Lu**, A Kummel and J Robertson, Passivating the sulfur vacancy in monolayer MoS_2. *APL Materials* 6, 066104 (2018).

[6] **H Lu**, Y Guo, H Li and J Robertson, Fermi level De-pinning for Dopant-free Silicon Solar Cells. Submitted to *Appl. Phys. Lett*. (2018)

[7] **H Lu**, Y Guo and J Robertson, passivate the grain boundary in 2D transition metal dichalcogenides, in preparation (2018)

[8] **H Lu**, Y Guo and J Robertson, doping effect to metal-insulator transition in VO_2, in preparation (2018)

[9] **H Lu**, Y Guo and J Robertson, Electronic structure of the crystal and amorphous carbon nitrides, in preparation (2018)

[10] **H Lu**, Y Zhai, R Pan and S Yang, An effective method of accelerating Bose gases using magnetic coils. *Chinese Physics B*, 9, p.033 (2014).

[11] R Pan., X Yue, X Xu, **H Lu** and X Zhou, Multiple photon-echo rephasing of coherent matter waves. *Physics Letters A*, 379, pp.691-695 (2015).

Conference Contribution

[1] **H Lu**, H Li and J Robertson, Si photovoltaic contacts – passivation or Fermi level unpinning, Workshop on Dielectrics in Microelectronics (WODIM), Berlin, Germany, June 2018.

[2] **H Lu** and J Robertson, Methods of passivating Sulfur vacancies in 2D MoS_2, Materials Research Society (MRS) spring meeting, Phoenix, US, April 2018.

[3] **H Lu**, Y Guo and J Robertson, Band Edge States, Defects and Dopants in Layered Semiconductors HfS_2 and SnS_2, Materials Research Society (MRS) spring meeting, Phoenix, US, April 2018.

[4] **H Lu** and J Robertson, Methods of passivating the sulfur vacancy in 2D MoS_2 48^{th} IEEE Semiconductor Interface Specialists Conference (SISC), San Diego, US, December, 2017.

[5] **H Lu**, Y Guo and J Robertson, Chemistry vs Dimensionality Effects in 2D semiconductor Contacts, E-MRS 2016 Fall Meeting, Warsaw, Poland, September, 2016.

Contents

Chapter 1 Introduction

1.1 Scaling the semiconductor devices

The Complementary metal-oxide-semiconductor devices, known as CMOS, are the primary technology for constructing integral circuits. CMOS has been widely applied in microprocessors, logic circuits and random access memories (RAM). CMOS consists of a pair of p-type and n-type field-effect transistors (FET), mostly metal-oxide-semiconductor FETs called MOSFETs. MOSFETs are the building block of every electronic device. Due to the increasing need for faster processing, higher integration and lower power consumption of devices, MOSFETs were scaled to smaller dimensions according to Moore's law in the past few decades. Moore's law says that the number of transistors in a dense integrated circuit doubles every 18 months.

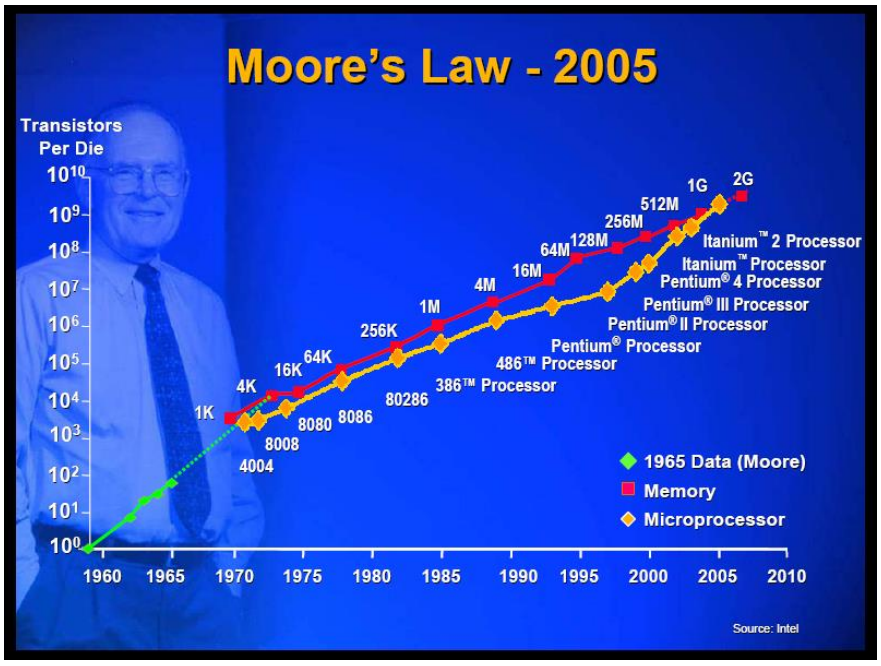


Fig 1-1. The Plot of CPU transistor counts against dates of introduction, from ref [1].

The simplest scaling concept is called constant-field scaling, which is scaling all dimensions, voltage, doping concentration and dielectric thickness. This results in the constant electric field, inversion-layer charge density, and carrier velocity and power density.

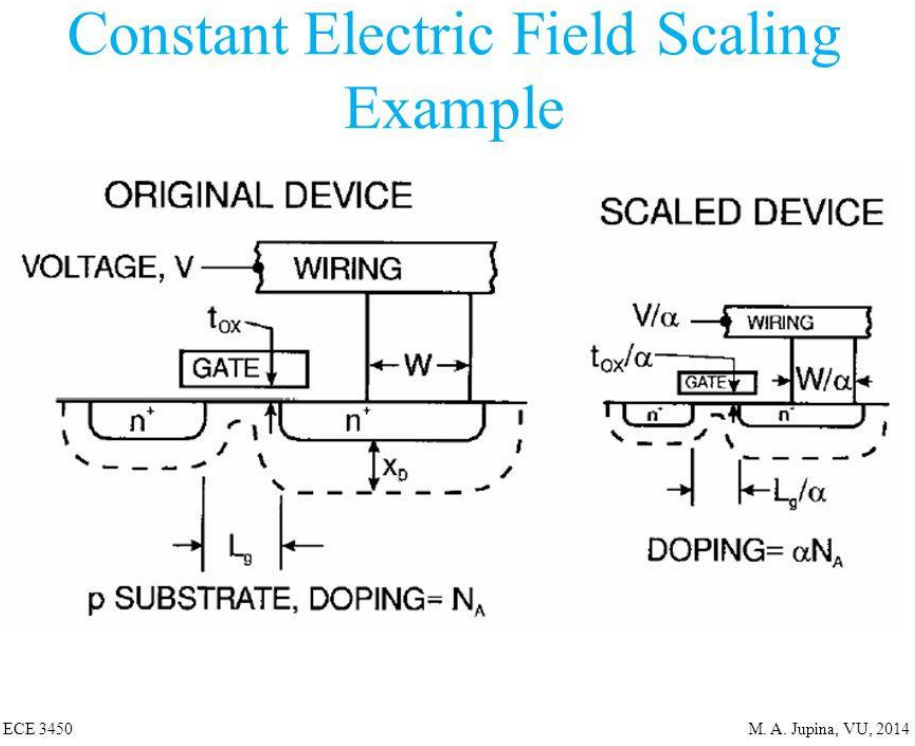


Fig 1-2. Constant field scaling of FET, from ref [2].

Si-based FETs are most widely used in today's electronic devices because of their low cost in massive fabrication and a good band gap of 1.1eV as the semiconductor for the channel. However, continuing scaling of Si devices is hard today as quantum effects become important as the gate length reduces below 10nm [3]. According to Zhirnov *et al.* [4], the minimal feature size of a "binary logic switch", based on Heisenberg uncertainty, is given by $x_{min} = \frac{\hbar}{\sqrt{2m_e E_S}} = \frac{\hbar}{\sqrt{2m_e k_B T ln2}} \approx 1.5nm$ at T=300K. Besides, the thickness of oxide should also be reduced with the device downscaling. For example, SiO_2 is the most commonly used gate dielectric. If its thickness reduces to less than a certain point, the tunnelling current to the channel is not acceptable. This certain point depends on the quality of SiO_2. An ultrathin SiO_2

layer with 1.4nm thickness has been reported to have low leakage current density [5]. Alternatively, high K dielectrics such as HfO_2, ZrO_2TiO_2, and Al_2O_3 have been introduced [6].

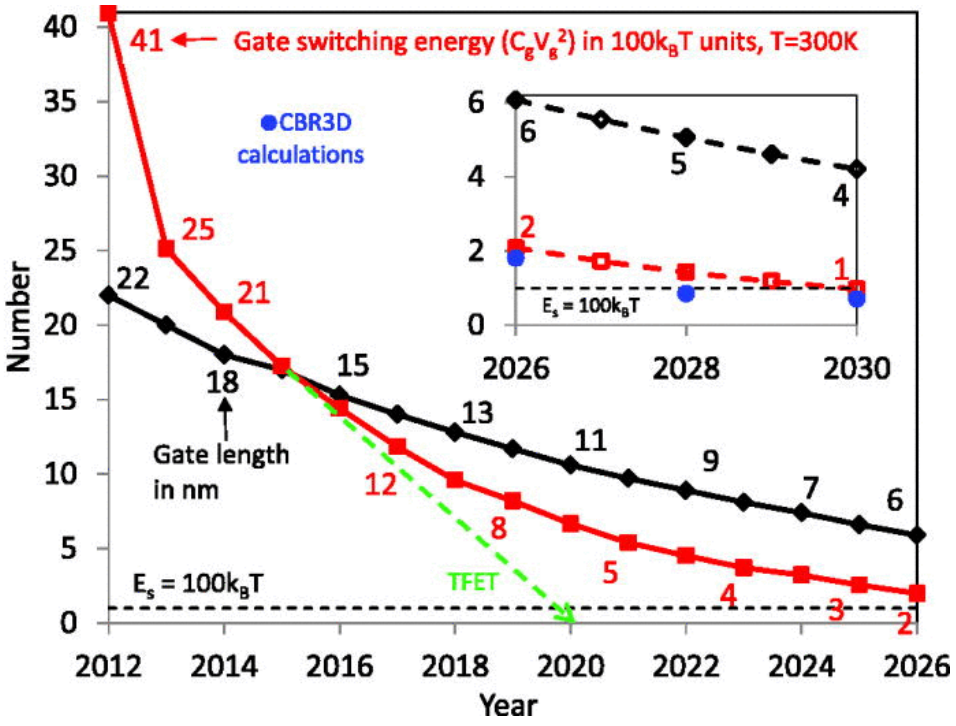


Fig 1-3. Scaling of the size of FET against year of introduction from ref [3].

The prediction of Moore will come to an end unless new materials are introduced to overcome the Si MOSFET's limits. 2D materials have attracted much attention as novel channel, electrode, and gate dielectric materials in the last five years. 2D materials include a large variety of ultra-thin materials from the gapless Graphene, semiconductor transition metal dichalcogenides (TMD) to Nitrides. Their bulk forms are all layer stacked materials. Instead of chemical bonds like in Si, the interlayer force of 2D materials is the weak van der Waals interaction [7]. Therefore in principle, there is no need to passivate the surface for defect-free 2D materials. From a device scaling point of view, the biggest advantage of 2D materials is their ultra-thin features, so that they suppress short-channel effects [8]. Therefore, dimensionally they perform better than Si in the 5nm gate length regime and below.

Graphene has no band gap so it is not considered as a channel material. However, it is a good electrode due to its high mobility [9]. TMD monolayers' band gaps lie within semiconductor range [10]. They are good alternatives to Si and III-Vs as the channel. Hexagonal boron nitride is a wide band gap material so it is not suitable for the channel. But it is a 2D high K material so it can be used as a gate dielectric. Apart from that, the stacking of 2D materials, called a van der Waals heterostructure, is useful to design band structures, called band engineering.

In this chapter, the most important 2D materials, graphene, TMDs and hexagonal nitrides will be introduced in order. Then we introduce the concept of different types of defects, which are the primary concern of studying the device performance. Finally, we illustrate the theory of Schottky barrier height (SBH) in a metal-semiconductor contact. It is important as it applies in electrodes connected to either gate, source or drain. Reducing the SBH is favoured in most cases so as to decrease the contact resistance and realize lower contact resistance and lower energy consumption.

1.2 2D materials from semi-metal to insulator

1.2.1 Graphene

Graphene is the thinnest material and the first '2D material' in the real sense. It is a 2D sheet without information in the third dimension. It is a hexagonal, carbon material which has relatively simple electronic structure compared with other 2D materials. The band structure near the Fermi level is in the shape of Dirac cone, where electron behavior is like massless Fermions, called Dirac Fermions. Therefore, its properties attract many physicists. For example, unusual half-integer Quantum Hall effects for both electrons and holes have been observed [11]. Under strain, Graphene's atomic and electronic structures are changed and the result can be described as an effective pseudo-magnetic field on the electronic degrees of freedom to induce pseudo-quantum Hall effect [12, 13].

Graphene is also highly valued for its excellent electronic and mechanical properties. It is a strong material with high Young's modulus as well as a high electron mobility up to 2 ×

$10^5 cm^2/(Vs)$ at room temperature [14]. The high mobility allows it to be applied in high speed electronic and optical devices, solar cells, durable display screens, and gas detection.

Graphene is a semimetal because it lacks a band gap, therefore it cannot be switched off in a transistor, and it cannot be the channel in FET in logic circuits. The research topic has been shifted from graphene as a single device to a combination with other 2D materials, for example, as the interlayer between semiconductor and metal to reduce the pinning, or just as an electrode. More attention is paid recently to other 2D materials like TMD, as they have band gaps so they can be used as the channel.

1.2.2 Transition metal dichalcogenides

Transition metal dichalcogenides (TMD) are a series of semi-conductive materials of the form MX_2 (M=Mo, W, Hf, Sn, etc.; X-S, Se, Te) stacked layer by layer via van der Waals force. They have a hexagonal unit cell. The monolayer MX_2 (M=Mo, W; X=S, Se, Te) all have direct band gaps in the semiconductor range while changing to indirect band gap in the bulk form [15], which means the monolayer can be used as the channel in FET devices and emitters in optical devices.

The strong in-plane bonding and weak inter-plane bonding leads to the fact that 2D TMDs can be obtained by mechanical exfoliation like Graphene, peeling from bulk by micromechanical cleavage using adhesive tape.

TMDs, from bulk to monolayers, break parity symmetry (which also happens in hexagonal nitride, but not Graphene), so there is no inversion center, which leads to nonlinear optical effect, such as second-harmonic generation. Apart from that, TMDs with 2H phase have direct band gaps in 2D form. The conduction and valence band edge is at the K point in the first Brillouin zone. Due to the loss of inversion center, the six K points shown in Fig. 4 are not equivalent. There are two different K point, namely K+ and K-, which introduces a new degree of freedom, valley polarization: the number of excitons in those two K valleys can be different by optical pumping. This degree of freedom can be used as a quantum bit, shown in Fig. 1-4 [16, 17].

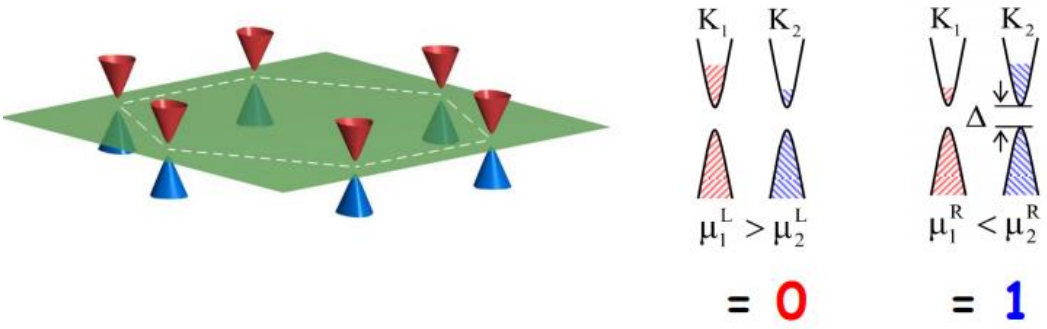


Fig 1-4. Control the number of excitons in valley via optical pumping and realize quantum bit, from ref [16].

Besides, valley symmetry leads to optical selection rules relying on the polarization of incident light, called spin-valley coupling. The spin-orbital coupling in heavy transition metals is more important at the band edge. For example in 2D MoS_2, d orbital splitting is about 0.15eV, shown in Table 1-1. The optical gap is different for the different spin direction, as shown in Fig 1-5.

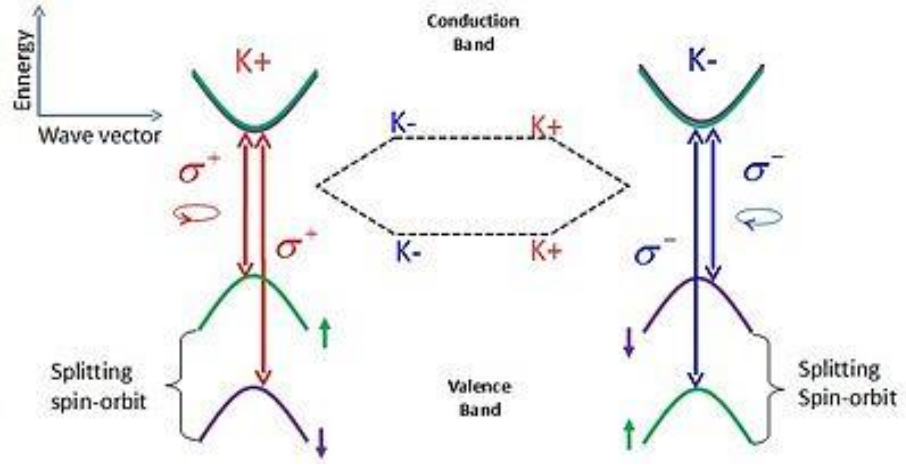


Fig 1-5. Spin splitting and optical selection rule, from ref [18].

Table 1-1. The calculated energy of the spin-orbit coupling, from ref [18, 19].

	Valence band splitting (eV)	Conduction band splitting (eV)
MoS_2	0.148	0.003
WS_2	0.430	0.026
$MoSe_2$	0.184	0.007
WSe_2	0.466	0.038
$MoTe_2$	0.219	0.034

The advantage of TMDs over graphene is that they have band gaps, although their mobility is not as high as graphene. They can be fabricated in a smaller size than conventional bulk semiconductors like Si, the on/off ratio of their FETs can be as high as 10^8 in MoS_2, although the mobility (410 $cm^2V^{-1}s^{-1}$, 300K) is a third of that in Si (1350 $cm^2V^{-1}s^{-1}$) [20].

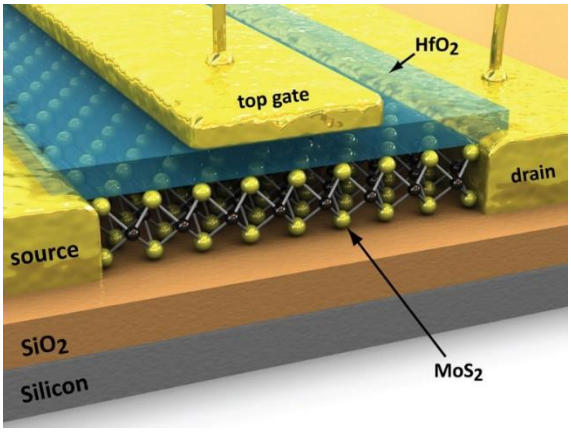


Fig 1-6. MoS_2 FET. Au as the source, gate, drain materials, HfO_2 as the gate dielectric, SiO_2 as the substrate, from ref [21].

1.2.2.1 MoS_2

Among all TMDs, MoS_2 is the most intensively studied one. There are three bulk crystal phases, 2H, 3R and 1T, all of which consist of the layer structure, held together by van der Waals bonding, shown in Fig 1-7. The monolayer has two crystal phases, 1T and 2H. 2H is more common and 1T is metastable and can be stabilized through doping with electron donors [22]. When heating with microwave radiation, it changes to the 2H phase [23]. The 1T phase can undergo a metal-insulator transition under uniaxial strain [24]. In this thesis, we will focus on 2H MoS_2, so all MoS_2 in this work refers to 2H MoS_2 unless specified.

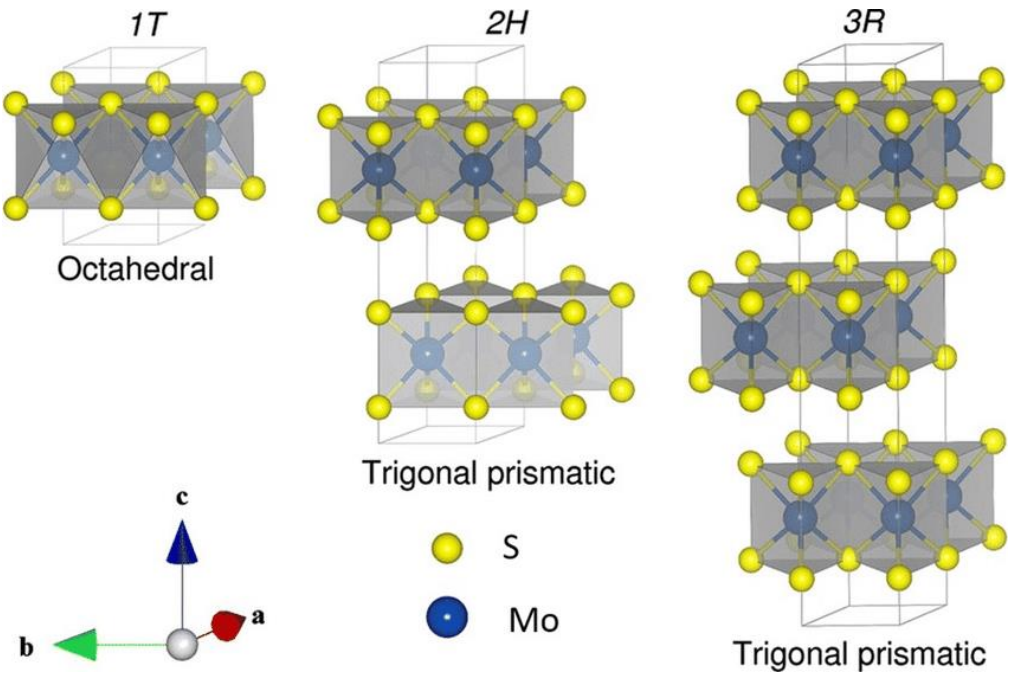


Fig 1-7. Three bulk phases of MoS_2. Ref [25]

Electronically, bulk MoS_2 has an indirect band gap of 1.29eV, when peeling it thinner and thinner to 2D, it shifts to a direct optical band gap of 1.90eV, revealed both by DFT calculation and experimental measurement [26, 27]. In monolayer form, the conduction band and valence band edge are at K in the first Brillouin zone, which makes it a good photoluminescence material, with orange light.

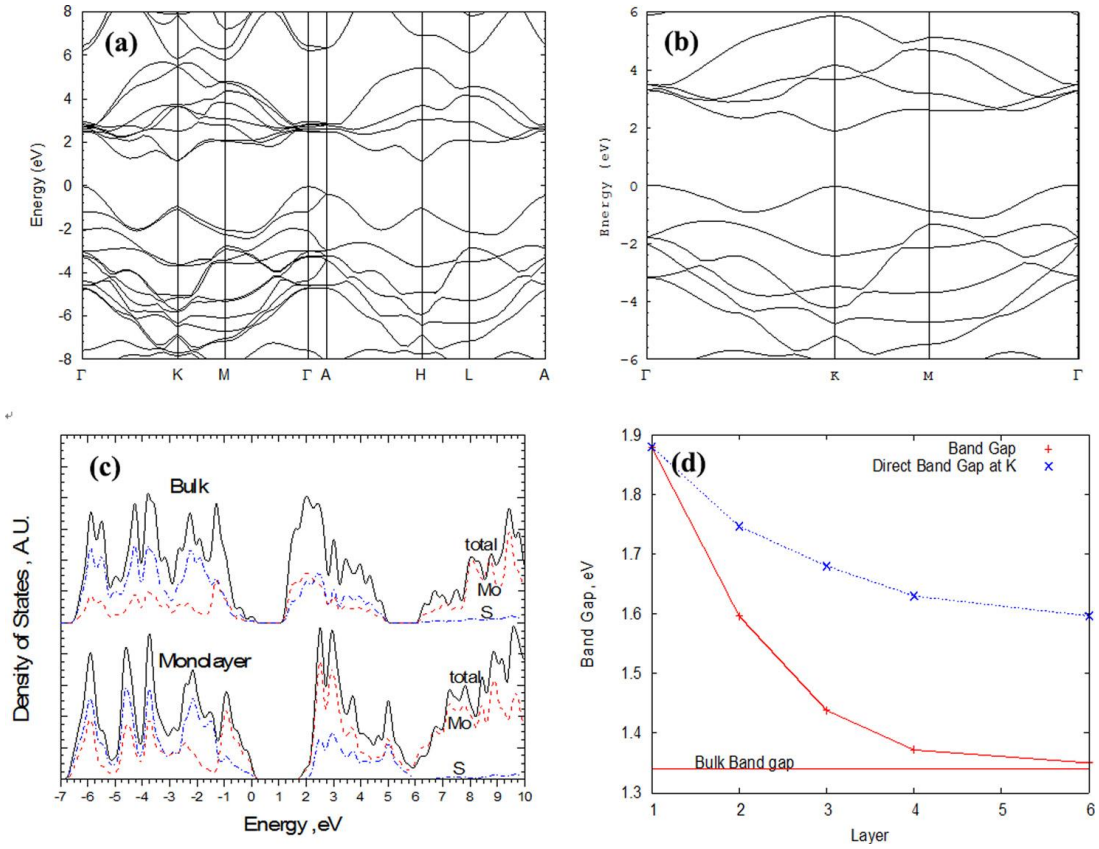


Fig 1-8. Band structure of (a) bulk (b) monolayer MoS_2 and Density of states (DOS) (c) of bulk and monolayer MoS_2. From ref [27].

As for the defects in 2D MoS_2, the sulphur vacancy has been proved to be the most significant intrinsic defect in the mechanical exfoliated product, while structural defects like the grain boundaries are prevalent in chemical vapour deposition (CVD) growth samples. The intrinsic mobility is not as high as Graphene, or even Si. With defects, the mobility is even lower. Therefore, many people try to passivate the defect using molecules, super acids (acids with acidities greater than that of 100% pure sulphuric acid [28]) or Cl. Another important application of defects in MoS_2 is in hydrogen evolution, including water splitting. It can serve as a catalyst for this process [29].

1.2.2.2 SnS_2 and HfS_2

We move on to another kind of TMD. Unlike MoS_2 which has 2H as the most stable phase, HfS_2 and SnS_2 are in the 1T phase. They are relatively cheap and the 2D form can be made via exfoliation using scotch tape [30, 31]. Their electronic structures are different from MoS_2 as well. They have indirect band gaps in 2D form. The d orbital is deep so that the valence bands and the conduction bands consist of s, p orbitals, which are less flat along the path in the Brillouin zone than the d orbital [32]. Therefore, the effective mass of electron and hole in band edge is small. For example, the effective mass of 2D HfS_2 is only 0.25 m* along the ΓK direction. More details of band structure are shown in chapter 6.

Their atomic structure is similar to that of 1T MoS_2 shown in Fig 1-7. Although HfS_2 and SnS_2 cannot be used in optical devices, they can be the channel in low energy FETs due to their high mobility compared to MoS_2 [30, 21]. For instance, they are good candidates for the n-type layers in the tunnel FET. Their large electron affinities make them available to fit with a p-type layers like WSe_2.

1.2.3 Hexagonal BN, AlN, and GaN

Boron Nitride (BN) exists abundantly with an amorphous form and various crystal forms, including hexagonal BN (h-BN), cubic-BN (c-BN) and wurtzite-BN (w-BN), shown in Fig 1-9. Cubic and wurtzite BN are by far the hardest materials except for diamond [33]. Hexagonal BN is the most stable phase and soft, and is often used in lubricants. The high electro-negativity of the B-N bond is strong, and bonding is covalent, making a hexagonal BN an insulator. Therefore, it is called white graphene. Bulk h-BN is a layered structure maintained by van der Waals force. Unlike c-BN and w-BN, the inhomogeneous structure of h-BN gives it weak out-of-plane stability and high in-plane stability, which is even flatter than Graphene. The flatness of h-BN makes it a good substrate contact with Graphene, which increases the Graphene mobility than suspend Graphene [34]. This is because it helps prevent Graphene from structure ripple due to the fact that suspend Graphene is not strictly 2D, which is predicted by Mermin-Wagner theorem: continuous symmetries cannot be spontaneously broken at finite temperature in systems with sufficiently short-range interaction in one and

two dimensions [35, 36]. Therefore, long-range fluctuations are favoured in energy and entropy, in suspend Graphene [37]. Although the electron conductivity of h-BN is low, the thermal conductivity can be as high as 2000W/(m·K) [38]. Therefore, h-BN can be an excellent gate dielectric, shown in Fig 1-10. Monolayer h-BN has a direct band gap in the region of ultraviolet light, thus it has potential application in LED devices.

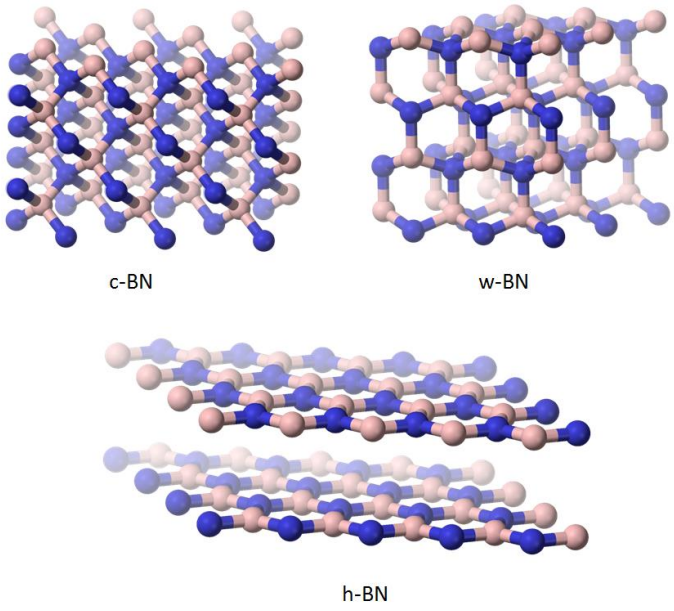


Fig 1-9. The three crystal phases of BN.

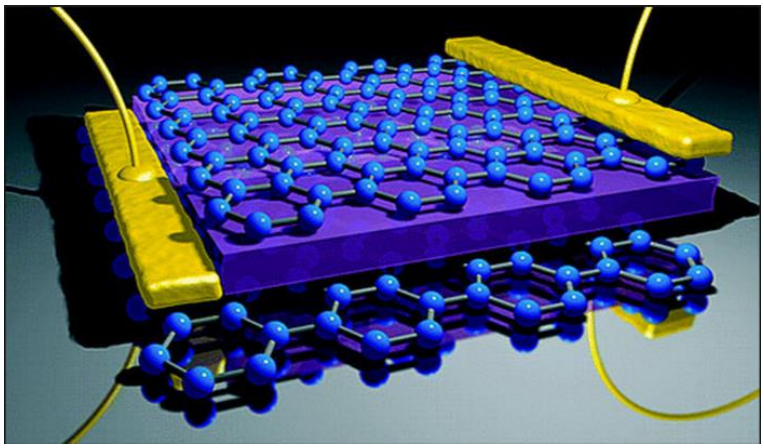


Fig 1-10. 2D h-BN as gate dielectrics, from ref [39].

In hexagonal XN (X=B, Al, Ga, In…), where the X moves from second element row downwards, the most stable phase changes from hexagonal to wurtzite. In a word, the X-N bond is more sp^3 like as the atomic number of X increases. The band gaps of monolayer h-XN are all direct and decrease from X=B to X=Al, Ga. The X-N bond in AlN and GaN is weaker than that of BN. Wurtzite AlN has piezoelectric properties so it can sense ultrasound [40]. Wurtzite GaN attracts more attention in optical electronics (LED) and solar cells [41]. The hexagonal form of AlN and GaN are much less studied and h-GaN has not been synthesized yet.

1.3 Defects

Defect formation is the most important matter if any material is considered to be industrialized and commercialized. No matter how good the pristine structure is predicted to be, it is not applicable if the spontaneous defects formed in nature or in the manufacture process can severely compromise the quality. The most common disadvantages of the defect are that it can act as a charge trap centre and scatter centre so lower the carrier mobility. It can also induce midgap states which can act as recombination centres in optical devices and lower the photoluminescence efficiency.

The defect is not always a bad thing. It makes Fermi level pinning stronger because it will increase the penetrating length of metal induced gap states, which can be used if you want to pin the Schottky barrier height in a fixed value.

The defects in most materials in the thesis can be classified according to their dimension. 0-dimensional defects affect isolated sites so they are called point defect. For example, a vacancy, a substitution, or an adatom. 1-dimensional defects are called dislocation, line defects, which is the broken pattern in crystal, seen abundantly in CVD grown 2D MoS_2 for example. The 2-dimensional defects are surfaces or grain boundaries. Since we study 2D materials, the grain boundary is 1D. The three-dimensional defects will occupy a finite volume in the crystal and of course, it is not periodical.

In this work, we will focus mainly on point defect and line defect, which are the major issue in most 2D materials.

1.3.1 Point defects

Point defects can be intrinsic or extrinsic, depending on whether foreign atoms are introduced or not. Fig 1-11 shows different types of point defects. Intrinsic defect, includes the vacancy, which is an atom missing from a position, the interstitial, which is when an atom occupies a site which originally was vacant, while an extrinsic defect is a substitutional atom which can also be viewed as doping, shown in Fig 1-11 (a). In crystal compound consisting of more than one species, there are more of types of point defects, denoted in Fig 1-11 (b).

In 3D materials, interstitials usually have high formation energy because they usually cause an unfavorable bonding. Vacancies, on the other hand, are prevalent. Therefore, a lot of materials are intrinsically n-type doped or p-type doped. In 2D materials, the vacancy is not always the major point defect because the interstitial can stay outside the plane rather than being squeezed into a small cell. For example, in 2D HfS_2 or SnS_2, the Sn and Hf interstitial are important as well, they can stay on the surface or form Hf-Hf and Sn-Sn bond in the plane. In MoS_2, however, the S vacancy is the most important point defect. Apart from the vacancy and interstitial, in compounds, a third type of intrinsic defect called the anti-site is abundant in materials whose atoms are weakly ionized like GaAs. For example, in Fig 1-11 (b), an As atom is replaced by Ga or a Ga atom replaced by As. Finally, a Frenkel pair consists of a pair of vacancy and interstitial, which is charge neutral globally but may induce local ionization.

The extrinsic point defects are foreign atoms such as in Fig 1-11 (b): replacing a Ga with In, called substitution, or as an interstitial like Boron in Fig 1-11 (b). Usually, small atoms like H, C, and B are found to be interstitial atoms and large atoms are to be substitutional atoms. Doping intentionally introduces extrinsic defects and is usually substitutional. It is an important method of controlling the carrier properties in semiconductor engineering. Doping can be classified as donors and acceptors, depending on the number of valence electrons the foreign atoms possess. For example, in Si n-type MOSFET, intrinsic Si can only conduct

when electrons are excited into the conduction band, therefore the conductivity is low. We want the major carrier to be electrons so we can replace some Si with P then an extra free electron exists and it can only occupy the conduction band. So this electron can move across the lattice freely. The conductivity of P doped Si is significantly enhanced. P is called donors. Alternatively, if we want p-doped Si, which causes a hole in the bonding pattern, we can replace some Si with B.

Point defects break the periodicity and lower crystal symmetry. Therefore, it induces gap states which are localized around the defect centre. In chapter 2, we will talk about how to calculate the formation energy of point defect with necessary correction of DFT error.

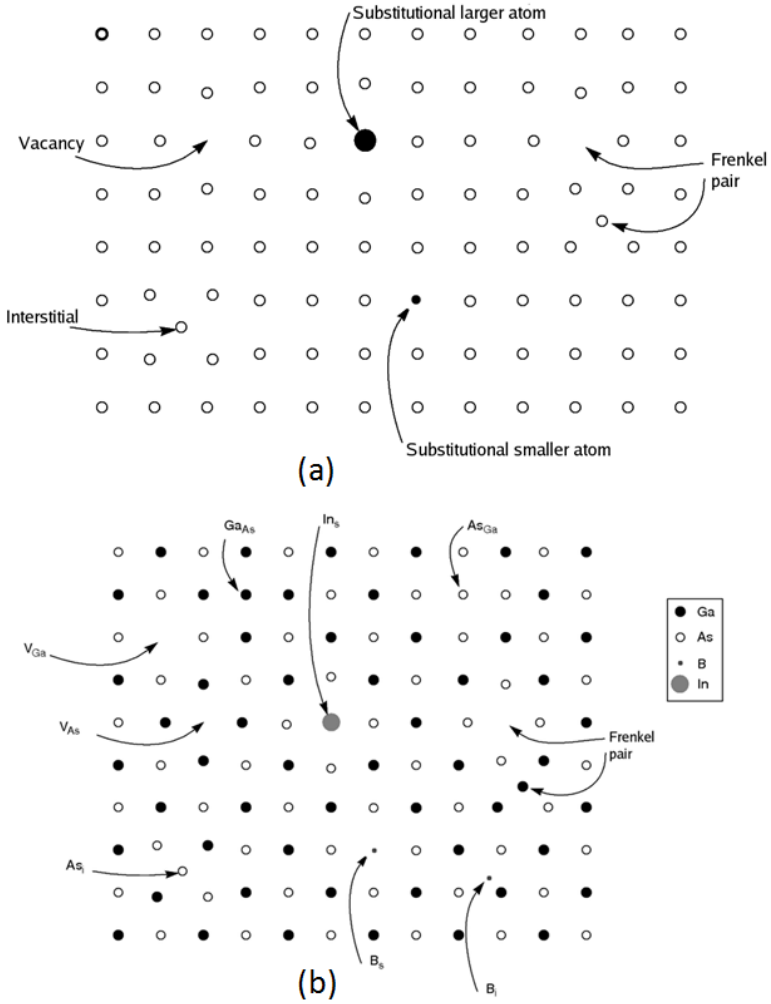


Fig 1-11. (a) different type of point defect in monoatomic crystal and (b) in the compound, using GaAs as an example, from ref [42].

1.3.2 Line defects

Like point defects, line defects break the periodicity as well and the lines can be very long so it is a global defect or structural defect, rather than point defect which is localized defect. Fig 1-12 shows the grain boundary of MoS_2 monolayer, the blue, yellow and red line is where the line defects are. Grain boundaries can also induce gap states, as shown in Fig 1-12 (d). The abundance of line defects in most CVD growth 2D materials like MoS_2 will damage the mobility even more than point defects.

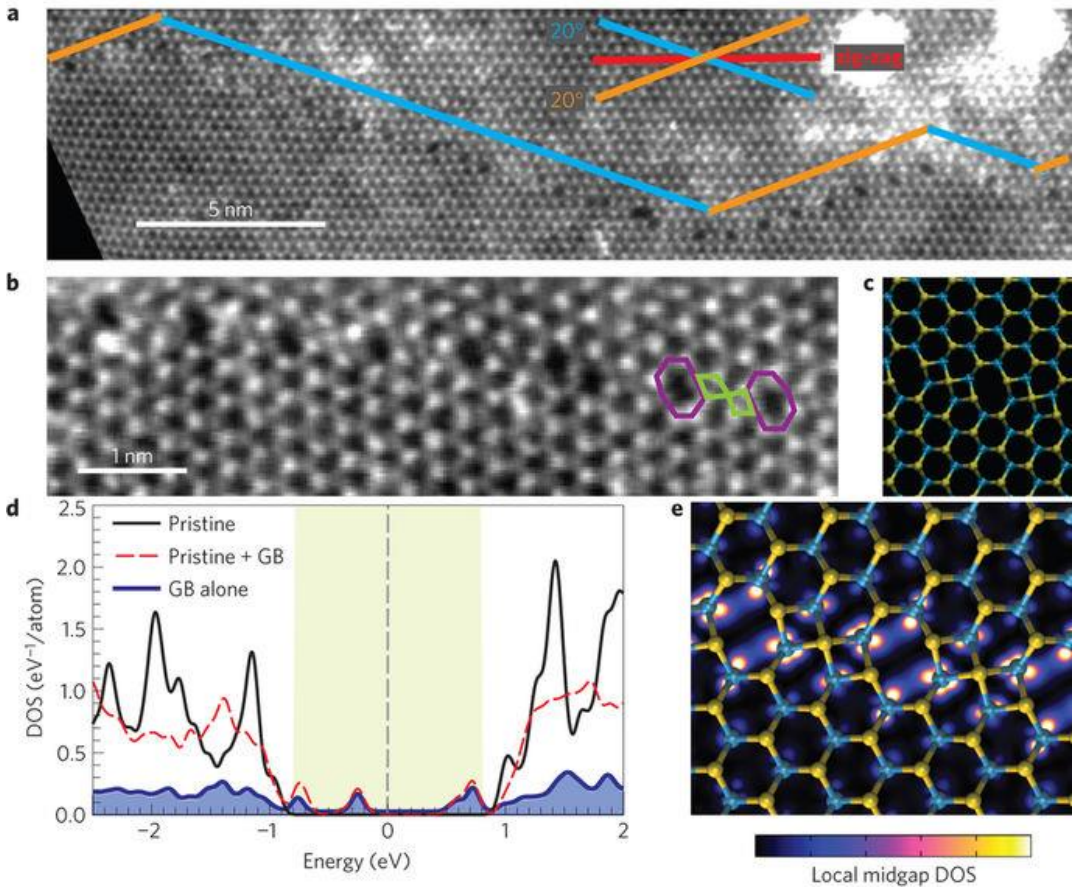


Fig 1-12. (a)(b) High-resolution ADF-STEM image of grain boundaries in 2D MoS_2. (c) Atomic model of the grain boundary shown in (b). (d) Total DOS of pristine, with grain boundary MoS_2. (e) A 2D spatial plot of the local mid-gap DOS, from ref [43].

1.3.3 Passivation

Where there is a defect, there may be a need to eliminate them or make them chemically and electronically inactive, called passivation. Passivation is defined as to remove all the defect states in the gap and return the Fermi level in the middle of the gap, like in a pristine semiconductor. Passivation is not a new concept, for example, H is a surface passivation agent in Si or diamond-like carbon to help reconstruct the surface [44]. Fig 1-13 shows the structure of Si surface passivated by hydrogen.

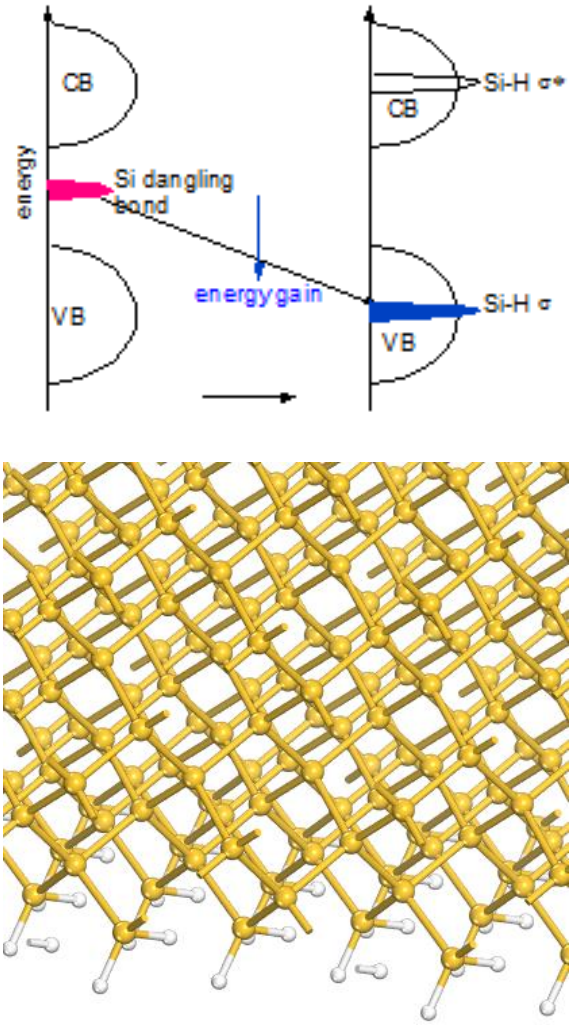
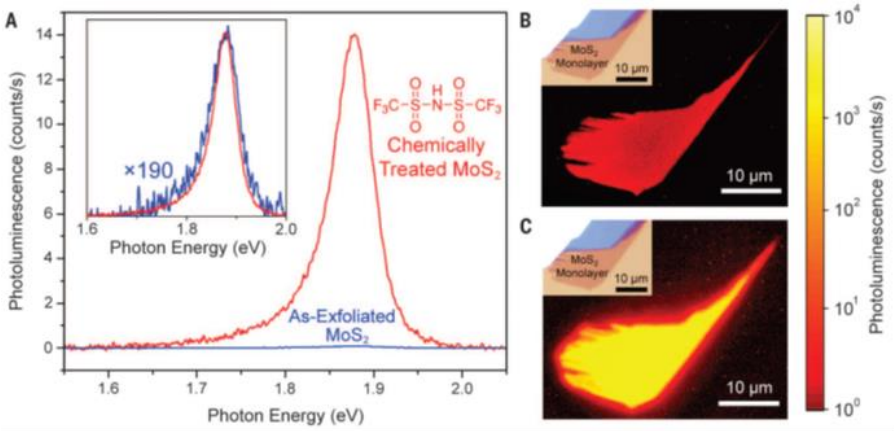


Fig 1-13. Hydrogen passivation of Si dangling bonds in Si (100) surface.

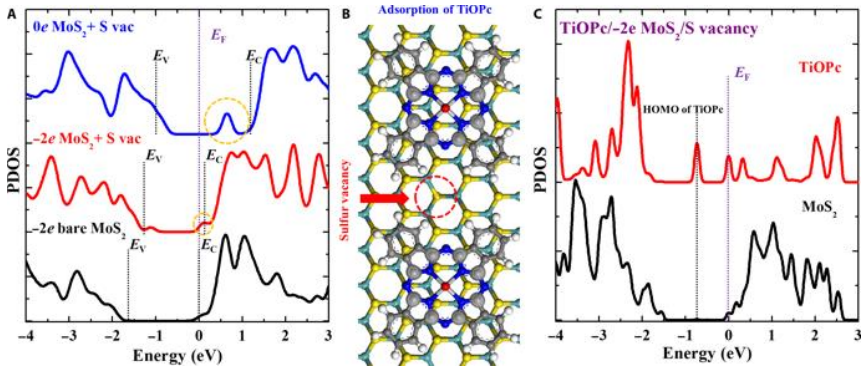
The unpaired electrons in the surface Si are in gap states shown red in Fig 1-13. Hydrogen forms the bonding state and anti-bonding state with surface Si atoms and separates them into the valence band and the conduction band. Therefore, all gap states are cleaned and electrons won't be trapped locally.

Passivation is not limited to surfaces, it can also fix point defects as long as the passivation agent can locate the defect centres and chemically treat them. It can be done by molecules, doping, surface contacts or even atoms and ions. The attraction of the passivation agent to the defect point is very important. It can be revised by changing the ambient environment. Fig 1-14 illustrates three examples of passivation of a sulphur vacancy in 2D MoS_2, which is by far the most popular 2D material. The first is using superacid to recover the photoluminescence (PL) to be 100 times higher [45]. The passivation agent is hydrogen. Details are shown in chapter 7. The second is using an organic layer to contact and clean gap states by charge transfer [46]. The third is a process by desulfurization of a thiol molecular fixing the S vacancy [47].

(a)



(b)



(c)

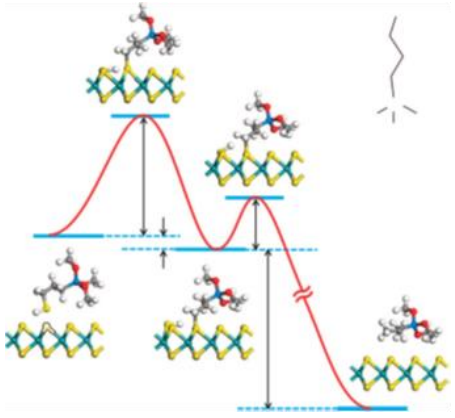


Fig 1—14. Three examples of passivating S vacancy in MoS_2, (a) super acid treatment, (b) organic layer, (c) thiol molecules, from ref [45, 46, 47].

The passivation of 2D materials is a hot topic since defects lowers device performance badly. In chapter 7, we will theoretically explain the passivation mechanism of the most popular 2D material, MoS_2.

1.4 Theory of Schottky Barrier Height

When electrons travel through an interface, there is a potential barrier scattering someelectrons are scattered back and some penetrate through. The interface lowers the carrier mobility and induces contact resistance. The device performance relies on not only the intrinsic mobility of semiconductor but also the contact resistance in the interface between metal electrodes and the semiconductor channel. Therefore, it is of great significance to investigate the properties at the interface. Usually, a low contact resistance is desirable in an FET so less energy is consumed, called the Ohmic contact where the current-voltage curve is linear, which means that there is no rectifying effect. In contrast, if the I-V relationship is not linear, then it is called Schottky contact. Whether it is Schottky contact or Ohmic contact depends on the Schottky barrier height (SBH). If the SBH is low or negative (Fermi level inside valence band or conduction band) then it is an Ohmic contact. Occasionally the Schottky contact is used as in Schottky diodes where the rectifying characteristic is preferred.

The n-type SBH or the electron barrier is defined as the difference between the Fermi level of the system and the CBM of the semiconductor. The p-type SBH or the hole barrier is defined as the difference between the Fermi level and the VBM of semiconductor which is same as subtracting the band gap of the semiconductor from the n-SBH. It is shown in Fig 1-15.

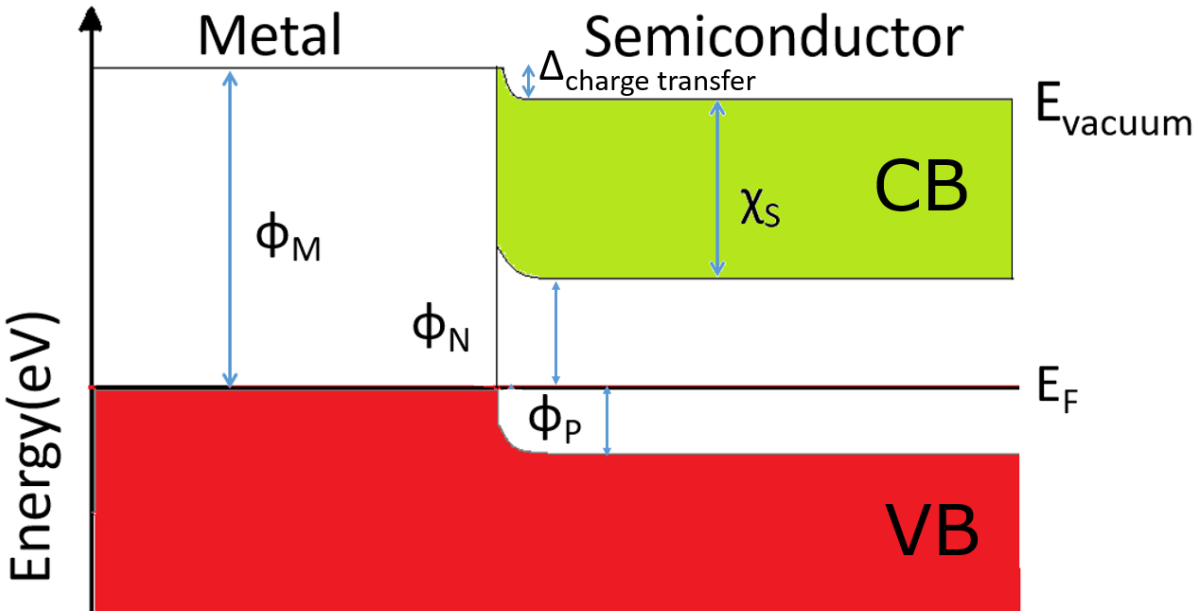


Fig 1-15. Band diagram of metal-semiconductor contact. ϕ_n is n-type SBH, ϕ_p is the p-type SBH, χ_S is electron affinity, ϕ_M is the metal work function.

The charge transfer in the interface determines the relation between SBH and metal work function. At the interface between to two metals of different work function, there is charge transfer induced potential to equalize the Fermi level of two metals. Similarly, at the interface between the metal and the semiconductor, there is charge transfer as well but not as strong as at the metal interface. The charge transfer induces a small potential step, as shown in Fig 1-15. In the Bardeen limit, the charge transfer is strong so that pins the Fermi level in a certain point inside band gap of the semiconductor. In the Schottky limit where there is no charge transfer, the n-SBH is described by the difference between the work function of the metal and electron affinity of the semiconductor.

$$\phi_n = \phi_M - \chi_S \tag{1.1}$$

Fermi level pinning (FLP) arises from Schottky limit to Bardeen limit, the charge transfer creates a dipole and thereby reduce the n-SBH to

$$\phi_n = \phi_M - \chi_S - \Delta \tag{1.2}$$

We are curious about how large is the charge transfer induced potential and we know that in the Bardeen limit where the SBH is pinned so it doesn't vary with metal work function. We then define that in strong pinning limit, the n-SBH is

$$\phi_n = \phi_{CNL} - \chi_S \tag{1.3}$$

CNL is the charge neutrality level which will be introduced shortly in this section. Right now it can be treated as the reference energy of Bardeen and Schottky limits. It is known that n-SBH varies linearly with metal work function. Then between these two limits, we can define the n-SBH as:

$$\phi_n = S(\phi_M - \phi_{CNL}) + (\phi_{CNL} - \chi_S) \tag{1.4}$$

S is the dimensionless Fermi level pinning factor between 0 (Bardeen limit) and 1 (Schottky limit). Then we know that the charge transfer potential is

$$\Delta = (1 - S)(\phi_M - \phi_{CNL}) \tag{1.5}$$

The origin of the charge transfer dipole and interface states which cause the FLP has been vividly discussed during last 40 years. The argument revolved around two principle models, the intrinsic models called metal induced gap states (MIGS) and the extrinsic model of defect states. Both of them can induce FLP.

1.4.1 Origin of Fermi level pinning

In the intrinsic case which is developed by Bardeen, Heine [48] and Flores [49], there are interface gap states induced by metal, due to the dangling bonds of semiconductor and metal wave states penetrating evanescently into the gap. These states are called MIGS or virtual gap states (VGS). The higher the density of MIGS are, the stronger fermi level pinning is. Fig 1-16 shows the metal induced gap states affecting Schottky Barrier Height.

MIGS are strong if there are chemical bonds between interfaces, in other words, if the metal is chemisorbed onto the semiconductor. One might think if there is no dangling bond in the interface, there is no MIGS, for example, in van der Waals interface where metal is physisorbed onto the semiconductor. This is incorrect because MIGS can still exist but they

decay faster over van der Waals distance. In chapter 3, a chemical trend of pinning strength related to adsorption condition in the interface is illustrated.

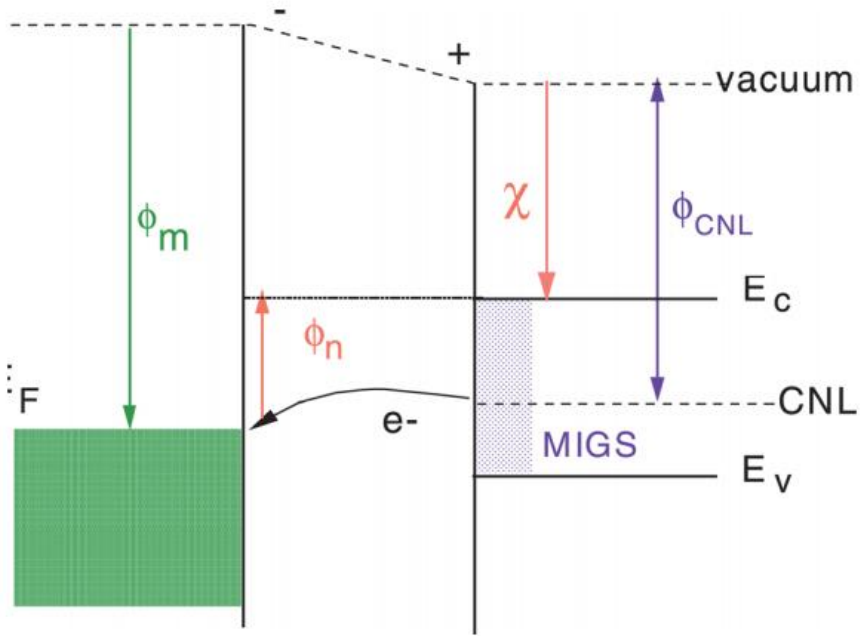


Fig 1-16. Schematic of a Schottky barrier showing charge transfer between metal and semiconductor and MIGS inside band gap, from ref [50].

Compared to MIGS, the extrinsic view seeks the fact about a chemical reaction in the interface which induces defect states and the degree of reaction strongly affects the SBH. This correlation is supported by the observation that the Fermi level on nonpolar (110) surfaces of III-V semiconductors is pinned by intrinsic defects and the SBH varies with the composition in the same way as the defect state energy [51].

In this work, the MIGS theory is preferred because there is a greater density of MIGS than defect states in most cases, so MIGS are more likely to cause pinning. Besides, MIGS explained the chemical trend of pinning factor S over a wide range of dielectric constants [52]. Finally, the theory of MIGS with CNL is completed and compatible with both experimental and Density Functional Theory (DFT) calculation where no defect occurs in between the interfaces.

1.4.2 Charge Neutrality Level (CNL)

We now explain what charge neutrality level (CNL) is and how it relates to the reference energy in the formula of SBH. The CNL is the branch point of the imaginary bulk band structure of the semiconductor in its band gap, where the band structure also contains evanescent interface states inside the gap and the k point is a complex number. It can be also understood as the energy above which the gap states are empty for a neutral surface [50]. The CNL can be derived as the solution of Green function:

$$G(E) = \int \frac{N(E')dE'}{E - E' + i\delta} = 0 \tag{1.6}$$

Where E is the energy level, N(E) is the density of states, δ is a small number to avoid singularity in the integral using residual theorem. The Green function integrates over the first Brillouin zone. It tells us that if we treat electron density of states as the source of a field in energy space, the CNL is the location where the field potential is zero. In DFT calculations, where energy level of even k point sampling over the first Brillouin zone, the formula can be expressed as a sum over all energy levels (the number of valence bands should be same as the number of conduction bands).

$$G(E) = \sum_i \frac{1}{E - E_i} = 0 \tag{1.7}$$

Then the CNL can be understood as a balance point of conduction band and valence band weight. For example, if the valence band density is high near VBM and conduction band density is low near CBM, then the CNL will be pushed near CBM, shown in Fig 1-17. The CNL is an intrinsic property of the bulk semiconductor which does not depend on the bonding of the interface or the metal it contacts with.

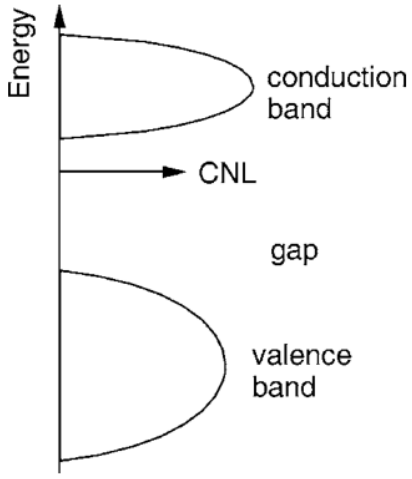


Fig 1-17. CNL is a weighted average point in the density of states (DOS). It is repelled by high density of states and attracted by flat bands.

1.4.3 The pinning factor S

The FLP factor S, which is the slope of the linear relationship between n-SBH and metal work function, is an intrinsic parameter of a semiconductor. Theoretically, it can be derived from equation (1.4) using the linear response model of Cowley and Sze [53] as

$$S = \frac{\partial \phi_n}{\partial \phi_M} = \frac{1}{1 + \frac{e^2 N \lambda}{\varepsilon}} \tag{1.8}$$

where e is the electronic charge, N is the density of the interface states per unit area, ε is the permittivity of semiconductor and λ is the decay length of interface states into the semiconductor. From the formula, we know that as the density of surface states increase, or the penetration depth increases, the pinning becomes stronger. On the other hand, high K materials usually resist pinning. Fig 1-18 shows that as the ionicity of semiconductor increases from non-ionic Si to high ionic SiO_2 and $SrTiO_3$, S rises dramatically.

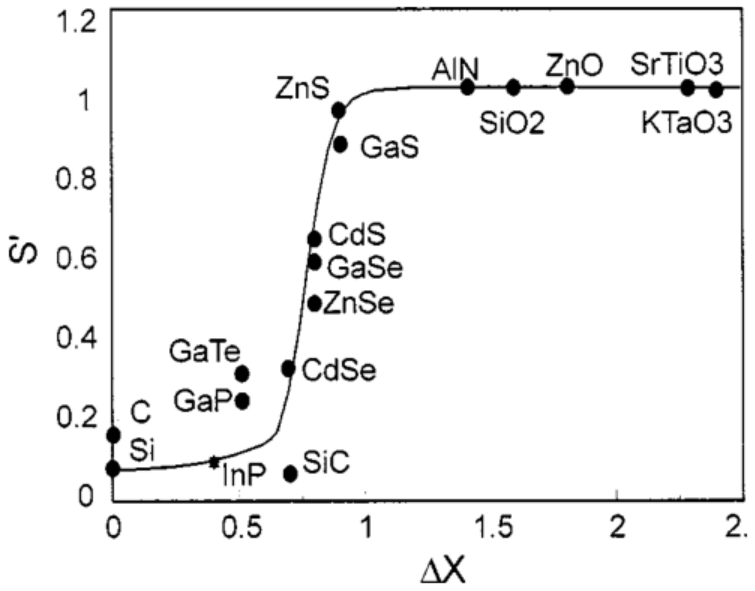


Fig 1-18. Pinning factor vs semiconductor electronegativity difference, from ref [54].

Apart from that, Mönch proposed that S depended on the optical dielectric constant [55]

$$S = \frac{1}{1 + 0.1(\varepsilon_\infty - 1)^2}$$

Because the calculated value of Nλ scaled with$(\varepsilon_\infty - 1)^{1.9}$, which is close to the power 2. It was initially believed that S varied with the semiconductor's ionicity. However, it is wrong [52]. For example, diamond and xenon have zero ionicity but small optical dielectric constant, so the S of them is large. Fig 1-19 plots the relationship of S and ε_∞, showing a linear dependence of log $(1\text{-}S^{-1})$ and log $(\varepsilon_\infty - 1)$.

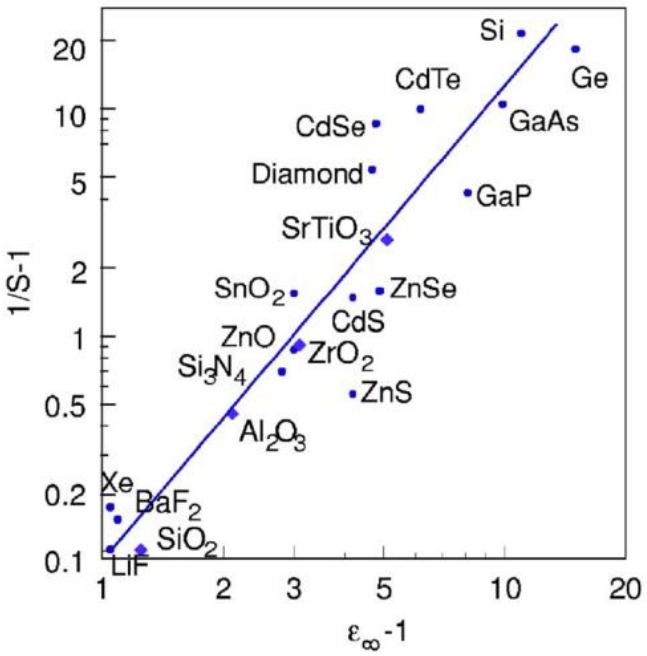


Fig 1-19. The plot of (S^{-1}-1) vs ε_∞-1 for various semiconductor, from ref [56].

1.5 Thesis aim and outline

The main aim of this thesis is to calculate the electronic characteristics like band structure, defect formation, defect passivation and doping of 2D materials as potential candidates for a next-generation semiconductor devices. The thesis outline is presented as follows:

Chapter 2 includes thorough details of the method used in this thesis, the Kohn-Sham scheme of density functional theory (DFT) and its several corrections like hybrid functional and dispersion correction. A correction of calculating charged defect is also provided.

Chapter 3 investigates the contact of metals with three 2D hexagonal Nitrides, BN, AlN, and GaN. A chemical trend is introduced to reveal the correlation between contact chemical environment and strength of Fermi level pinning.

Chapter 4 adds an oxide layer to tune the band offset and Fermi level pinning factor of Si. It concludes that the number of layers of the HfO_2 and the polarization of oxide interface will change the pinning factor and whole SBH offset, respectively.

Chapter 5 shows a new way of doping called transfer doping compared to conventional substitutional doping, which creates defects and lowers the mobility of graphene. Transfer doping can shift the Fermi level of graphene away from the Dirac point, while not breaking the surface.

Chapter 6 presents the intrinsic defect formation energy of HfS_2 and SnS_2, which are the n-type building blocks of tunnel FET (TFET), forming type II band alignment with WSe_2. HfS_2 and SnS_2 have low effective mass and therefore have higher mobility than MoS_2. The major type of defects includes S vacancies, S interstitials, and Sn/Hf interstitials.

Chapter 7 calculates several passivation schemes of S vacancies in 2D MoS_2 including substitutional doping, O_2, and superacid. We find H passivation in superacid is the most successful way to clean all gap states and return the Fermi level to midgap. Symmetry plays an important role in the chemical adsorption of H onto the S vacancy and we find the most energetically favorable and symmetrical configuration 3H -1 charge, passivates the S vacancy.

Chapter 8 provides the conclusions of this work and the perspectives of future work.

Reference

1. https://machprinciple.com/post/50-years-of-Moore-s-law-at-Silicon-Valley
2. Frank, David J., et al. Proceedings of the IEEE **89** 259 (2001).
3. M Denis, and X Gao. Appl. Phys. Letters **106** 193503 (2015).
4. V V Zhirnov, R K Cavin III, J A Hutchby, and G I Bourianoff, Proc. IEEE **91** 1934 (2003).
5. Asuha, et al. Appl. Phys. Lett. **81** 3410 (2002).
6. J Robertson, Rep. Prog. Phys. **69** 327 (2005).
7. W Zhang, Q Wang, Y Chen, Z Wang and A T S Wee, 2D Materials, **3** 2 (2016).
8. R Granzner, Z Geng, W Kinberger and F Schwierz. (2016, October). 2016 13th IEEE International Conference on (pp. 466-469). IEEE.
9. R Garg, S Elmas, T Nann and M R Andersson, Adv. Energy Mater. **7** 1601393 (2017).
10. J Gusakova, X Wang, L L Shiau, A Krivosheeva, V Shaposhnikov, V Borisenko, V Gusakov and B K Tay, Phys. Status Solidi A, **214** 1700218 (2017).
11. Y Zhang, Y W Tan, H L Stormer and P Kim, Nature, **438** 201 (2005).
12. A C Neto, F Guinea, N M Peres, K S Novoselov and A K Geim, Rev. Mod. Phys, **81** 109 (2009).
13. M R Masir, D Moldovan and F M Peeters, Solid State Communication **175-176** 76 (2013).
14. K I Bolotin., et al. Solid State Communications **146** 351 (2008).
15. Y Guo, D Liu and J Robertson, ACS Appl. Mater & Inter. **7** 25709 (2015).
16. A M Jones, et al. Nat. Nanotech. **8** 634 (2013).
17. H Zeng, et al. Nat. Nanotech. **7** 490 (2012).
18. G B Liu, W Y Shan, Y Yao, W Yao and D Xiao, Phys. Rev. B. **88** 085433 (2013).
19. Y C Cheng, Z Y Zhu, M Tahir and U Schwingenschlogl, EPL. **102** 57001 (2012).
20. Q H Wang, et al. Nat. Nanotech. **7** 699 (2012).
21. B Radisavljevic, A Radenovic, J Brivio, V Giacometti and A Kis, Nat. Nanotech. **6** 147 (2011).
22. A N Enyashin., et al. J. Phys. Chem. C **115** 24586 (2011).
23. D Xu, et al. Nanotechnology **27** 385604 (2016).
24. X Lin and N Jun, Phys. Lett. A **379** 2883 (2015).
25. B Han and H Yun, Energy Science & Engineering **4** 285 (2016).
26. K F Mak, et al. Phys Rev. Lett. **105** 136805 (2016).
27. D Liu, Y Guo L Fang and J Robertson. Appl. Phys. Lett. **103** 183113 (2013).
28. N F Hall and J B Conant, J. Am. Chem. Soc. **49** 3047 (1927).
29. G Ye, et al. Nano Lett **16** 1097 (2016).
30. T Kanazawa, T Amemiya, A Ishikawa, V Upadhyaya, K Tsuruta, T Tanaka and Y Miyamota, Sci. Rep. **6** 22277 (2016).
31. D Chu, S W Pak and E K Kim, Sci. Rep. **8** 10585 (2018).
32. H Lu, Y Guo and J Robertson, Appl. Phys. Lett. **112** 062105 (2018).
33. G Leichtfried et al. (2002). "13.5 Properties of diamond and cubic boron nitride". In P. Beiss; et al. Landolt-Börnstein – Group VIII Advanced Materials and Technologies: Powder Metallurgy Data. Refractory, Hard and Intermetallic Materials. 2A2. Berlin: Springer. pp. 118–139

34. E V Castro, et al. Phys. Rev. Lett. **105** 266601 (2010).
35. P C Hohenberg, Phys. Rev. **158** 383 (1967).
36. N D Mermin, H Wagner, Phys. Rev. Lett. **17** 1133 (1966).
37. A O'Hare, F V Kusmartsev and K I Kugel, Nano Lett. **12** 1045 (2012).
38. I Jo, et al. Nano Lett. **13** 550 (2013).
39. L Britnell, et al. Nano Lett. **12** 1707 (2012).
40. S Walter, et al. "Investigations on aluminum nitride thin film properties and design considerations for smart high frequency ultrasound sensors." The Second International Conference on Smart Sensor, Actuators and MEMS, Prague, Czech Republic. 2011.
41. C Y Chen, G Zhu, Y Hu, J Yu, J Song, K Cheng, L Peng, L Chou and Z Wang ACS Nano **6** 5687 (2012).
42. H Hausmann, A Pillukat and P Ehrhart, Phys. Rev. B **54** 8527 (1996).
43. V Zande, M Arend, et al. Nat. Mater. **12** 554 (2013).
44. D B Fenner, D K Biegelsen and R D Bringans, J. Appl. Phys. **66** 419 (1989).
45. M Amani, et al. Science **350** 1065 (2015).
46. J Park, et al. Science Advances **3** e1701661 (2017).
47. Z Yu, et al. Nat. Comm. **5** 5290 (2014).
48. V Heine, Phys. Rev. **138** 1689 (1965).
49. C Tejedor, F Flores and E Louis. J. Phys. C: Solid State Physics **10** 2163 (1977).
50. J Robertson, J. Vac. Sci. & Tech. A: Vacuum, Surfaces, and Films **31** 050821 (2013).
51. W E Spicer, et al. Thin Solid Films **89** 447 (1982).
52. J Robertson, J. Vac. Sci. & Tech. B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena, **18** 1785 (2000).
53. A M Cowley and S M Sze. J. Appl. Phys. **36** 3212 (1965).
54. J Tersoff, Phys. Rev. B **30** 4874 (1984).
55. W Mönch, "Role of virtual gap states and defects in metal-semiconductor contacts." Electronic Structure of Metal-Semiconductor Contacts. Springer, Dordrecht, 1990. 224-227.
56. W Mönch, Surface science **299** 928 (1994).

Chapter 2 Methods

In this chapter, we give a brief review of the methods used in the thesis, the theory of density functional (DFT), based on the fundamental theory of quantum physics. The spirit of DFT is to solve a many-body Schrödinger equation numerically and derive physical properties out of it, with the aid of some approximations which vary in different systems. The object of these approximations is to simplify the problem, enhance the calculation speed while keeping essential physical properties. The chapter provides a general framework of DFT or ab-initio calculation and the practice of approximation, step by step. At the end of this chapter, two corrections are introduced.

2.1 From many-body Schrödinger equation to Hartree-Fock method

Once quantum theory was built, the effort to utilize it to calculate the properties of materials in microscopic view has never ended. DFT is the way towards those goals. However, many approximation methods are needed. For example, if considering the relativity effect of electrons, one can try to solve the Dirac equation. However, the amount of electrons we need to calculate is huge, as many as ~10^{23}. The Dirac equation cannot be solved elegantly in analytic form and even worse as the system grows large it is not possible to make a numerical solution due to the lack of computer power. The Dirac equation, in most cases, can be simplified to the Schrödinger equation in the low energy limit in most cases. In some heavy transition metal systems like MoS_2 or WS_2, the effect of special relativity can be calculated as a perturbation as well, namely fine structure, including a Darwin term, a spin-orbital correlation term and a kinetic relativistic term. In addition, we only solve the time-independent system in this thesis. There are conditions where time-dependent DFT (TDDFT) is necessary, like exciton formation, GW method, etc. However, we will introduce the basic

DFT technics, which is to calculate the static ground state properties. Therefore, we start from the N electrons, S ions-body Schrödinger equation:

$$\boldsymbol{H}\psi(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N,\boldsymbol{R}_1,\dots,\boldsymbol{R}_S) = E\psi(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N,\boldsymbol{R}_1,\dots,\boldsymbol{R}_S) \tag{2.1}$$

where,

$$\boldsymbol{H} = -\sum_{\boldsymbol{i}}\frac{\hbar^2}{2m}\nabla_{\boldsymbol{r}_i}^2 + \frac{1}{4\pi\varepsilon_0}\sum_{\boldsymbol{i<j}}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} - \frac{1}{4\pi\varepsilon_0}\sum_{\boldsymbol{i,j}}\frac{Z_j e^2}{|\boldsymbol{r}_i-\boldsymbol{R}_j|} - \sum_{\boldsymbol{j}}\frac{\hbar^2}{2M_j}\nabla_{\boldsymbol{R}_j}^2 + \frac{1}{4\pi\varepsilon_0}\sum_{\boldsymbol{i<j}}\frac{Z_i Z_j e^2}{|\boldsymbol{R}_i-\boldsymbol{R}_j|} \tag{2.2}$$

In the equation, $\mathbf{r}_i$ is the i^{th} electron's spatial position and $\mathbf{R}_j$ is the j^{th} ion's spatial position. M_j is the mass of the j^{th} ion, m is the mass of an electron, which is closed to electron rest mass. The Hamiltonian contains the kinetic energy of both ions and electrons, as well as the Coulombic potential energy of electron-electron interaction, electron-ion interaction and ion-ion interaction. The first simplification we do is to treat ions as completely static although ions can vibrate around a series of fixed points in the lattice, weakly compared to electrons because the mass of an electron is a thousand times smaller than the mass of a proton and thus it is negligible. The approximation is called Born-Oppenheimer, also as an adiabatic approximation [1]. The Hamiltonian is then

$$\boldsymbol{H} = -\sum_{\boldsymbol{i}}\frac{\hbar^2}{2m}\nabla_{\boldsymbol{r}_i}^2 + \frac{1}{4\pi\varepsilon_0}\sum_{\boldsymbol{i<j}}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} - \frac{1}{4\pi\varepsilon_0}\sum_{\boldsymbol{i,j}}\frac{Z_j e^2}{|\boldsymbol{r}_i-\boldsymbol{R}_j|} \tag{2.3}$$

We know that the electron has spin and obeys the Pauli Exclusion Principle, therefore, the N-electrons wave function can be assumed as

$$\psi(\boldsymbol{r}_1,s_1;\dots;\boldsymbol{r}_N,s_N) = \frac{1}{\sqrt{N!}}\det\left[\psi_i(\boldsymbol{r}_j,s_i)\right] \tag{2.4}$$

$\psi_i(\boldsymbol{r}_j,s_j)$ is the wave function of a single electron. The Hamiltonian consists of two parts, one is the one-body operator and the other is the two-body operator. These operators are not always commutable to wave functions, of course.

$$\boldsymbol{H} = \sum_{\boldsymbol{i}} \boldsymbol{h}(\boldsymbol{r}_i) + \sum_{\boldsymbol{i<j}} \boldsymbol{g}(\boldsymbol{r}_i, \boldsymbol{r}_j) \tag{2.5}$$

Now we calculate the energy which is$< \psi|\boldsymbol{H}|\psi >$. For the one-body operator **h**, we can simplify the many body problem to one body

$$< \psi \left| \sum_{\boldsymbol{i}} \boldsymbol{h}(\boldsymbol{r}_i) \right| \psi >$$

$$= \frac{1}{N!} \int d\boldsymbol{r}_1 \dots d\boldsymbol{r}_N \sum_P (-1)^s P_s[\psi_P^*(\boldsymbol{r}_P, s_P)] \sum_{\boldsymbol{i}} \boldsymbol{h}(\boldsymbol{r}_i) \sum_Q (-1)^t Q_t[\psi_Q(\boldsymbol{r}_Q, s_Q)]$$

$$= \sum_{\boldsymbol{i}} \int d\boldsymbol{r}_i \, \psi_i^*(\boldsymbol{r}_i, s_i) \boldsymbol{h}(\boldsymbol{r}_i) \psi_i(\boldsymbol{r}_i, s_i) = \sum_{\boldsymbol{i}} \int d\boldsymbol{r} \, \psi_i^*(\boldsymbol{r}, s_i) \boldsymbol{h}(\boldsymbol{r}) \psi_i(\boldsymbol{r}, s_i) \tag{2.6}$$

We can split the spin wave function and orbital wave function and derive that the energy of one-body energy is

$$= \sum_{\boldsymbol{i}} \int d\boldsymbol{r} \, \psi_i^*(\boldsymbol{r}) \boldsymbol{h}(\boldsymbol{r}) \psi_i(\boldsymbol{r}) \tag{2.7}$$

The P and Q is the sum of all permutations of N electron single wave functions. The two-body operator **g** which is the electron-electron interaction part.

$$< \psi \left| \sum_{\boldsymbol{i>j}} \boldsymbol{g}(\boldsymbol{r}_i, \boldsymbol{r}_j) \right| \psi >= \frac{1}{N!} \int d\boldsymbol{r}_1 \dots d\boldsymbol{r}_N \sum_P (-1)^s P_s[\psi_P^*(\boldsymbol{r}_P, s_P)] \sum_{\boldsymbol{i<j}} \boldsymbol{g}(\boldsymbol{r}_i, \boldsymbol{r}_j) \sum_Q (-1)^t [\psi_Q(\boldsymbol{r}_Q, s_Q)]$$

$$= \frac{e^2}{8\pi\varepsilon_0 N!} \int d\boldsymbol{r}_1 \dots d\boldsymbol{r}_N \sum_P (-1)^s P_s[\psi_P^*(\boldsymbol{r}_P, s_P)] \sum_{\boldsymbol{i \neq j}} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \sum_Q (-1)^t [\psi_Q(\boldsymbol{r}_Q, s_Q)]$$

$$= \frac{e^2}{8\pi\varepsilon_0} \int d\boldsymbol{r} d\boldsymbol{r}' \sum_{\boldsymbol{i \neq j}} \frac{\psi_i^*(\boldsymbol{r}, s_i)\psi_j^*(\boldsymbol{r}', s_j)\psi_i(\boldsymbol{r}, s_i)\psi_j(\boldsymbol{r}', s_j) - \psi_i^*(\boldsymbol{r}, s_i)\psi_j^*(\boldsymbol{r}', s_j)\psi_i(\boldsymbol{r}', s_i)\psi_j(\boldsymbol{r}, s_j)}{|\boldsymbol{r} - \boldsymbol{r}'|}$$

$$= \frac{e^2}{8\pi\varepsilon_0} \int d\boldsymbol{r} d\boldsymbol{r}' \sum_{\boldsymbol{i \neq j}} \frac{|\psi_i(\boldsymbol{r})|^2 |\psi_j(\boldsymbol{r}')|^2 - \psi_i^*(\boldsymbol{r})\psi_j^*(\boldsymbol{r}')\psi_i(\boldsymbol{r}')\psi_j(\boldsymbol{r})\delta_{s_i,s_j}}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{2.8}$$

The one-body operator **h** does not only consist of electron-ion interaction and kinetic energy of electrons, but may also contain another external field, we just define the external

field interaction to the electrons V_{ex}, so the Schrödinger equation of the i^{th} electron can be written as

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{ex}(\boldsymbol{r}) + \frac{e^2}{8\pi\varepsilon_0}\sum_{j\neq i}\int\frac{|\psi_j(\boldsymbol{r}')|^2}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}'\right]\psi_i(\boldsymbol{r}) - \left[\frac{e^2}{8\pi\varepsilon_0}\sum_{j\neq 1}\int\frac{\psi_j^*(\boldsymbol{r}')\psi_i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\delta_{s_i,s_j}\psi_j(\boldsymbol{r})d\boldsymbol{r}'\right] = \varepsilon_i\psi_i(\boldsymbol{r}) \quad (2.9)$$

This is called **Hartree-Fock equations**, which is non-linear [2]. The first term contains the electron kinetic energy, electron-electron interaction and electron interaction with external fields. However, the second term is pure quantum mechanics which is called the exchange term. This term is only non-zero when two coupling electrons have the same spin. Due to the Pauli Exclusion Principle, electrons with the same spin do not like to be too close to each other. As a result, each electron has a 'hole' associated with it known as an exchange hole or a Fermi hole. The charge of the 'hole' is positive so is equivalent to the absence of electron around it.

The limitation of the Hartree-Fock approximation is that it assumes the N-body wave function of the system can be well-represented by a single Slater-determinant [3]. **The single-determinant approximation does not take into account Coulomb correlation and treats each electron's wave function independently.** The exact wave function is not accessible with the Hartree-Fock approach. The energy difference of the real non-relativistic ground state energy and the Hartree-Fock energy, called the correlation energy is generally negative, shown in Fig 2-1.

$$E_C = E_0 - E_{HF} \quad (2.10)$$

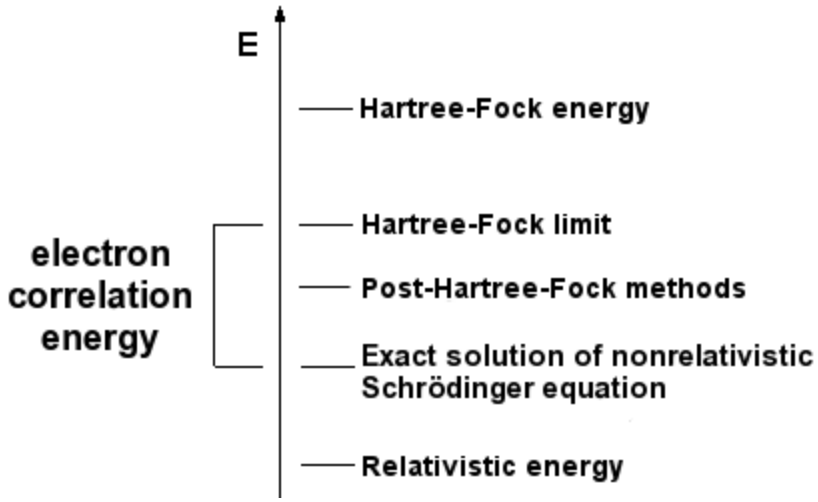


Fig 2-1. Electron correlation energy in terms of the various levels of theory of solutions for the Schrödinger equation, from ref [2].

2.2 Density Functional Theory

Since the wave function of an electron is not an observable and we don't care about which electron is label 1, 2, 3…N because the electron is indistinguishable. A physical observable property is the electron density in a spatial distribution, which is the summation over all individual electrons with spin:

$$n(\boldsymbol{r}) = N \sum_{s_1 \dots s_N} \int d\boldsymbol{r}_2 \dots d\boldsymbol{r}_N |\psi(\boldsymbol{r}, s_1; \dots; \boldsymbol{r}_N, s_N)|^2 = 2 \sum_{\boldsymbol{i}} \psi_i^*(\boldsymbol{r}) \psi_i(\boldsymbol{r}) \qquad (2.11)$$

Using electron density as the variable rather than each electron's wave function helps us construct the one-electron Schrödinger-like equation of a fictitious system called Kohn-Sham system which consists of non-interacting electrons that generate the same density as any given system of interacting electrons [4]. This simplification saves a lot of resources and makes the calculation of large system possible. We want to make the total energy a functional of total electron density n(**r**). From the Hartree-Fock equation, we try to interpret the total energy with electron density:

$$E[n(\boldsymbol{r})] = -\frac{\hbar^2}{2m}\int 2\sum_{\boldsymbol{i}} \psi_i^*(\boldsymbol{r})\nabla^2\psi_i(\boldsymbol{r})\,d\boldsymbol{r} + \int n(\boldsymbol{r})V_{ex}(\boldsymbol{r})d\boldsymbol{r} + \frac{e^2}{8\pi\varepsilon_0}\int \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}d\boldsymbol{r}' + E_{XC}[n(\boldsymbol{r})] \quad (2.12)$$

The total energy has four parts, the kinetic energy part, which cannot be written explicitly as a functional of electron density; the external field parts and Hartree energy of electron-electron Coulomb energy; the exchange and correlation energy which includes the exchange term in the Hartree-Fock equation and Coulomb correlation energy which cannot be expressed by Hartree-Fock equation. We can separate each term and define Kohn-Sham Hamiltonian as:

$$\boldsymbol{H} = -\frac{\hbar^2}{2m}\nabla^2 + V_{eff}(\boldsymbol{r}) \quad (2.13)$$

$$V_{eff}(\boldsymbol{r}) = V_{ex}(\boldsymbol{r}) + V_H(\boldsymbol{r}) + V_{XC}(\boldsymbol{r}) \quad (2.14)$$

$$V_H(\boldsymbol{r}) = \frac{e^2}{4\pi\varepsilon_0}\int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{d}\boldsymbol{r}' \quad (2.15)$$

$$V_{XC}(\boldsymbol{r}) = \frac{\delta E_{XC}[n(\boldsymbol{r})]}{\delta n(\boldsymbol{r})} \quad (2.16)$$

Where effective potential consists of external field potential V_{ex}, electron-electron interaction V_H and exchange-correlation term V_{XC}. We notice that only the minimum value of Kohn-Sham energy has physical meaning. There are two important theorems about the ground state energy and electron density:

Theorem I: For any system of interacting particles in an external potential, the potential is determined uniquely by the ground state particle density n(**r**), except for a constant shift.

Theorem II: A universal functional for the total energy E[n(**r**)] in terms of n(**r**) can be defined, valid for any external potential. For any particular external field, the exact ground state energy of the system is the global minimum value of this functional, and the density n(**r**) corresponding to the minimum energy functional is the exact ground state density.

The functional alone is sufficient to derive any ground state physical properties. However, excited states of electrons cannot be determined by the energy functional. The Kohn-Sham equations represent a mapping of the interacting many-electron system onto a system of non-interacting electrons moving in an effective potential screened by other electrons. An iterative step is used to solve the Kohn-Sham equations:

1. Define an initial, trial electron density n(**r**).
2. Solve the Kohn-Sham equations using the trial electron density to find the single-particle wave functions, $\psi_i(\boldsymbol{r})$.
3. Calculate the n(**r**) using

$$n(\boldsymbol{r}) = 2\sum_{\boldsymbol{i}} \psi_i^*(\boldsymbol{r})\psi_i(\boldsymbol{r}) \tag{2.17}$$

4. Compare the calculated electron density with the trial one. If the two densities' difference is within the threshold, then this is the ground state electron density and it can be used to calculate the total ground state energy. If not, the trial n(**r**) must be updated to the calculated one and repeat step 2 and 3 until reaching the converge criteria.

Above is called self-consistent field (SCF) procedure, which is a standard procedure of energy minimization. To get the minimum of the Kohn-Sham functional, method called steepest descents is applied. With the information of only the energy functional, the direction of next step can be obtained via the negative of a gradient operator $-\frac{\partial E}{\partial n}$ acting on the vector of electron density [5]. When the gradient goes close to zero, the minimization is converged and finished.

2.3 Periodic system and plane wave basis set

Right now we have obtained a single-body Hamiltonian and energy functional to sufficiently determine the ground state energy and density. However, the equation is not linear, the wave function is a field of the infinite degrees of freedom, it extends over the entire solid, the basis set required to expand each wave function is infinite. The differential

equations with integrals are hard to solve numerically. Besides, the expression of exchange and correlation functional is far from tractable. The difficulties can be all surmounted by expanding the wave function by a plane wave basis set, if the system, as well as the wave function, is periodic like, for example, in crystal structure.

Bloch theorem states that in a periodic solid each electronic wave function can be written as the product of a cell-periodic part and a wavelike part [6].

$$\psi_i(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_i(\boldsymbol{r}) \tag{2.18}$$

The cell-periodic part of the wave function can be expanded with a discrete basis set of plane waves.

$$u_i(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{i,\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \tag{2.19}$$

Due to the periodicity of $u_i(r)$, the reciprocal lattice vectors **G** obeys

$$\boldsymbol{G}\cdot\boldsymbol{l} = 2\pi m; m \in \mathbb{Z} \tag{2.20}$$

where **l** is the lattice vector of the periodical cell. Then each electron wave function can be expanded by a plane wave set as

$$\psi_i(\boldsymbol{r}) = \sum_{\boldsymbol{G}} c_{i,\boldsymbol{k}+\boldsymbol{G}} e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}} \tag{2.21}$$

2.4 K-point sampling and energy cut-off

In a cell, the number of electrons is finite and gives rise to energy levels but the selection of k points is infinite. Thus the Bloch theorem changes the problem of calculating an infinite number of wave functions to calculating a finite number of wave functions at an infinite

number of k points. The k points can be selected from the first Brillouin zone because any point outside it has an equivalent k point in the first Brillouin zone.

The Kohn-Sham equation in k space is then:

$$\sum_{\boldsymbol{G'}}[\frac{\hbar^2}{2m}|\boldsymbol{k}+\boldsymbol{G}|^2\delta_{\boldsymbol{G},\boldsymbol{G'}}+V_{eff}(\boldsymbol{G}-\boldsymbol{G'})]c_{i,\boldsymbol{k}+\boldsymbol{G'}}=\varepsilon_i c_{i,\boldsymbol{k}+\boldsymbol{G}} \tag{2.22}$$

The equation is a linear equation rather than a differential equation. To solve the equation in any given k point, we need diagonalization of the Hamiltonian matrix whose dimension depends on the choice of cut-off energy. The size of the matrix can be reduced if using pseudopotential method in section 2.5..

The choice of k point selection is infinite so it is not approachable. Fortunately, the solutions of two k space Kohn-Sham equation are nearly identical if those two k points are close to each other. In other word, electrons with similar momentum may have the same collection of eigenvalues and eigenstates. Therefore, we can represent the wave function over a small region of k space by the wave function at a single k point. Hence we only need to sample a finite number of k points in the first Brillouin zone to obtain the energy level of a periodic system. A method of sampling k points in the first Brillouin zone has been developed to obtain an accurate approximation to the electronic potential and the contribution to the total energy from a filled electronic band, called Monkhorst-Pack method [7]. The spirit of this method is to sample the k points as uniformly as possible. An example is shown in Fig 2-2 of how to sample k points uniformly in the 2D hexagonal cell.

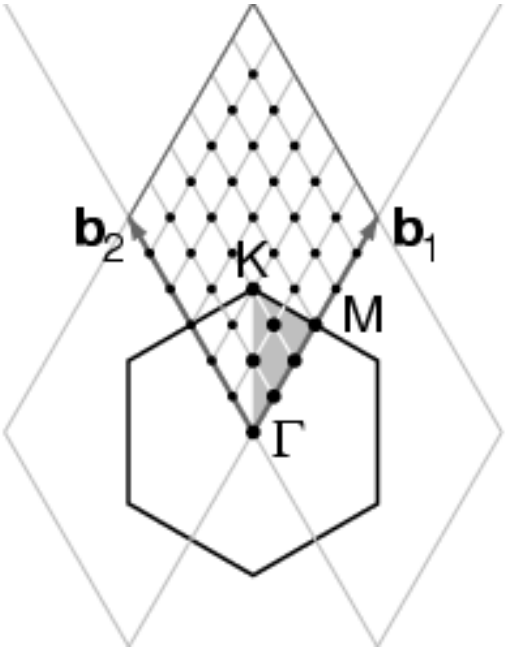


Fig 2-2. Monkhorst-Pack method of sampling k points in the 2D hexagonal lattice.

With k point sampling, one can obtain a nearly correct electronic potential and total energy of an insulator or semiconductor by calculating a small number of k points. As for the metal, a denser k point is needed to find the Fermi surface. Any error induced by insufficient k point sampling can be corrected by using a denser set of k points. As more k points are put into the calculation, the total energy will converge.

After choosing the value of k, we now move to choose the vector G. The summation of all G vectors is not possible because there is an infinite number of G vectors fulfilling$\boldsymbol{G} \cdot \boldsymbol{l} = 2\pi m$. Therefore, we need to select the most important G vectors, whose coefficient $c_{i,\boldsymbol{k}+\boldsymbol{G}}$ is large. Usually, an electron wave function's high frequency part is negligible, which means electrons are less likely to have very high momentum. Therefore, we only consider the electrons whose kinetic energies are smaller than a certain energy called cut-off energy.

$$\frac{\hbar^2}{2m}|\boldsymbol{k}+\boldsymbol{G}|^2 \leq E_{cutoff} \tag{2.23}$$

The truncation of the plane-wave basis will lead to computation error if the cut-off energy is too low. The error can be reduced by increasing the cut-off energy and the suitable cut-off energy depends on the system and is generally tricky. A standard protocol is to do a total

energy converge test with cut-off energy from small to large (k point from sparse to dense of course) to find the proper value which is not too large to bring about unrealistic computational cost while not too small to relax into unrealistic structure and render the wrong total energy and wave function.

It is noteworthy that the method of expanding the wave function with a plane wave basis set, sampling k point, and truncating large kinetic energy can also be used in a non-periodic system such as defect, surface and isolated molecular. The key is to set a large supercell so that in the small region it is non-periodic but forms a large periodic system. If there is a vacuum, you need to set the length of vacuum large enough so the electrostatic potential along the direction attenuates to zero. If there is a defect, you need to construct big enough cell to isolate the defect from its periodic mirror images so that they are not mutually affected.

2.5 Pseudopotential

In the last section, we mentioned that we set a cut-off energy so the plane wave sets whose the kinetic energies are larger than it are discarded. However, the effective potential field induced wavefunction usually contains high-frequency parts especially in deep levels close to the nucleus. Examples are the orbitals in a hydrogen-like atom, the wave function of s orbital of valence electron oscillates dramatically in the core region due to the Pauli Exclusive Principle, shown in Fig. 3 [2]. Therefore, the cut-off energy should be very high to include the high-frequency term and perform the all-electron calculation, which is too expensive.

The deep level electron with high-frequency terms is strongly localized around the core, which is less interested in material properties. On the contrary, the valence electron which is non-local determines most physical properties of solid, to a much greater extent. Therefore, we can revise the potential and its corresponding wave function with fewer high-frequency terms. The spirit is to set a cut-off core radius r_c, outside of which the new potential looks exactly like the all-electron potential and gives the same scattering properties, the new wave function looks exactly like the original one as well. Inside the r_c, the core electrons have been removed and the potential is revised to be weaker and smoother, so the potential and wave

function can be expanded with a small number of the small k-vector plane waves. We call the new potential and wave function pseudopotential and pseudo-wave function. They should be designed carefully so that outside cut-off radius they are identical to all-electron potential and wave function, shown in Fig 2-3. To ensure that, the pseudo-wave function should fulfill

$$\int_0^{r_c} |\psi_{pseudo}|^2 d\boldsymbol{r} = \int_0^{r_c} |\psi_{real}|^2 d\boldsymbol{r} \tag{2.24}$$

This requirement is called **norm-conserving**.

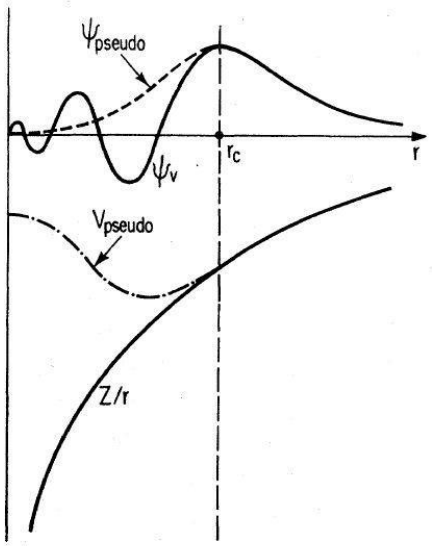


Fig 2-3. Illustration of all-electron (solid line) and pseudo-electron (dashed line) potentials and their corresponding wave functions. The radius at which all-electron and pseudo-electron values match is designated r_c, from ref [2].

Fig 2-4 illustrates the typical procedure of generating an ionic pseudopotential for an atom. It is a nontrivial process. In general, pseudopotential depends on the angular momentum of the state and the general form can be written as

$$V_{NL} = \sum_{lm} |lm > V_l < lm| \tag{2.25}$$

where |lm> are the spherical harmonics and V_l is the pseudopotential for angular momentum l. Those pseudopotentials whose V_l remains constant are called local pseudopotential, which is a function only of the distance from the nucleus.

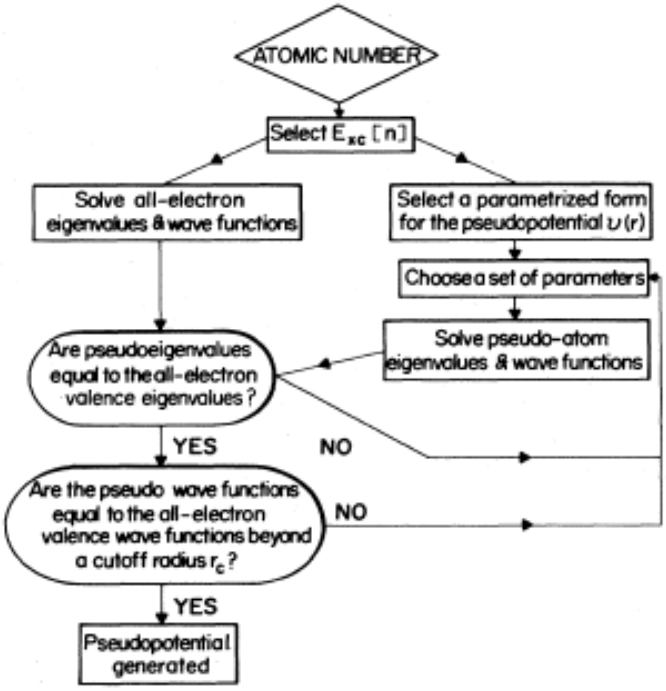


Fig 2-4. A flow chart describing the construction of an ionic pseudopotential for an atom, from ref [5].

It is known that there is an inherent limitation on optimizing the convergence of norm-conserving pseudopotentials. In order to lower the energy cut-off for the plane wave basis set, Vanderbilt introduces ultra-soft pseudopotential (USP) [8]. In most cases, a high cut-off energy is only required for the plane-wave basis set when there are tightly bound orbitals that have a substantial fraction of their weight inside r_c. Therefore, the only way to reduce the cut-off energy is to construct a potential violating the norm-conserving condition by removing the charge associated with these orbitals from the core region. The pseudopotential can then be as soft as possible within the core and thus lower the energy cut-off.

2.6 Exchange-Correlation Functional

The Kohn-Sham effective potential consists of the Hartree energy, the external field, and the exchange-correlation energy. Among those terms, only the exchange-correlation energy functional cannot be written in an analytic form. It cannot be expressed even as an explicit functional of the electron density. Nevertheless, we can define some general properties. The electron-electron interaction and exchange-correlation term together can be written as a 2-body functional:

$$E_H + E_{XC} = \frac{e^2}{8\pi\varepsilon_0}\int \frac{P(\boldsymbol{r},\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}d\boldsymbol{r}' \tag{2.26}$$

where

$$P(\boldsymbol{r},\boldsymbol{r}') = n(\boldsymbol{r})n(\boldsymbol{r}') + n(\boldsymbol{r})n_{XC}(\boldsymbol{r},\boldsymbol{r}') \tag{2.27}$$

Where n is the electron density. The first term gives rise to the Hartree energy and the second term is the exchange-correlation term. Classically, P(**r**,**r'**) can be interpreted as the probability of finding an electron in **r** and finding another in **r**'. Therefore,

$$\int P(\boldsymbol{r},\boldsymbol{r}')d\boldsymbol{r}' = (N-1)n(\boldsymbol{r}) \tag{2.28}$$

$$\int P(\boldsymbol{r},\boldsymbol{r}')d\boldsymbol{r}'d\boldsymbol{r}' = (N-1)N \tag{2.29}$$

In section 2.2 we said that the exchange term can be treated as a hole in the vicinity of an electron which is caused by the screening effect of other electrons. We can write the exchange and correlation term separately,

$$n_{XC}(\boldsymbol{r},\boldsymbol{r}') = n_X(\boldsymbol{r},\boldsymbol{r}') + n_C(\boldsymbol{r},\boldsymbol{r}') \tag{2.30}$$

and define

$$\int n_X(\boldsymbol{r},\boldsymbol{r}')d\boldsymbol{r}' = -1 \tag{2.31}$$

$$\int n_C(\boldsymbol{r}, \boldsymbol{r}')d\boldsymbol{r}' = 0 \tag{2.32}$$

2.6.1 LDA and GGA

To express the exchange-correlation energy as a functional of electron density, we need an approximation. The simplest approximation is called local-density approximation (LDA). In LDA, we assume that the exchange-correlation energy per electron at point **r**, which is $\varepsilon_{XC}(\boldsymbol{r})$, is equal to the exchange-correlation energy per electron in a homogeneous electron gas that has the same density as n(**r**).

$$\varepsilon_{XC}^{LDA}(\boldsymbol{r}) = \varepsilon_{XC}^{hom}[n(\boldsymbol{r})] \tag{2.33}$$

LDA assumes that the exchange-correlation energy is purely local and ignores the nearby inhomogeneity of the electron density. It is proven to work really well for non-magnetic materials. LDA results in an overestimation of binding between atoms and therefore an underestimation of bond length and lattice constant. Since we said that the LDA neglects the different density effect of exchange-correlation energy in the vicinity of a point, we can introduce the generalized gradient approximation (GGA) which not only depends on local density but also its first gradient.

$$\varepsilon_{XC}^{GGA}(\boldsymbol{r}) = \varepsilon_{XC}^{hom}[n(\boldsymbol{r})]F_{XC}[n(\boldsymbol{r}), \nabla n(\boldsymbol{r})] \tag{2.34}$$

The most common functional used in GGA was developed by Perdew, Burke, and Ernzerhof, called PBE [9]. GGA fixes the error of bond and lattice so it is comparably cheap and good to calculate the atomic structure and the electronic structure, where the correlation effect is weak. However, it still ignores the non-local effect. As a result, the band gap it calculates is always been underestimated.

2.6.2 DFT+U

In a magnetic system with strong correlation effect, like a Mott insulator, as an example, electrons are strongly localized and experience strong Coulomb repulsion. The many-body characteristics become significant. LDA and GGA tend to have over-delocalized valence

electrons and over-stabilized metallic ground state. We need a correction of the exchange-correlation term.

One of the ways to rationalize the physics of strong-correlated systems is the Hubbard model whose real-space second-quantization formalism is ideally suited to describe systems with electrons localized on atomic orbitals, although it is semi-quantitative and semi-empirical. The Hubbard Hamiltonian is

$$\boldsymbol{H}_{Hub} = T \sum_{<i,j>,s} (c_{i,s}^{\dagger} c_{j,s} + h.c.) + U \sum_{i} n_{i,\uparrow} n_{i,\downarrow} \tag{2.35}$$

where <i, j> denotes nearest-neighbor atomic sites, $c_{i,s}^{\dagger}$ is electron creation operator, $c_{j,s}$ is annihilation operator and $n_{i,\downarrow}$ is the number operator. The T is kinetic energy term with external field, U is term of electron-electron interaction. In strongly-localized systems, the electron's motion is described by a 'hopping' process from one site to its nearest-neighbour whose amplitude T is proportional to the dispersion of valence electronic states and represents the single-particle term of total energy. The strong Coulomb repulsion only exists in electrons on the same atom in the same state, with a different spin. The U is more empirical. In general, the strength of U and T determines whether the system is insulating or conducting. When T>>U, the hopping between localized sites is strong enough to overcome Coulomb repulsion from other electrons on neighbour sites, so the system is metallic. When T<<U, the system is an insulator.

The total energy functional in DFT+U is

$$E_{DFT+U}[n(\boldsymbol{r})] = E_{DFT}[n(\boldsymbol{r})] + E_{Hub}\left[n_{mm'}^{ls}\right] - E_{dc}[n^{ls}] \tag{2.36}$$

To avoid double-counting of the interaction energy both in E_{DFT} and E_{Hub}, we need to subtract a double-counting term E_{dc}. In double-counting term, a coefficient called J also needs to set up like U and the effective U_{eff} is

$$U_{eff} = U - J \tag{2.37}$$

As for how to obtain the empirical U and J, this can be computed them from the linear response, see ref [10]. DFT+U works well in open-shell systems with strongly localized

electrons such as transition metal oxides. However, most semiconductors are closed-shell and DFT+U gives limited improvement for the band gap correction. Therefore, a different functional with the non-local component is needed.

2.6.3 Hybrid functional

There is a functional beyond GGA which has non-local terms included. The Hartree-Fock (HF) method contains exchange term but no correlation term and the electrons are over localized and the band gap is always overestimated. The hybrid functional mixes Hartree-Fock potential with local functional like GGA to give right electron localization. Generally, the degree of mixing varies with the different functionals and can be interpreted as follows

$$E_{XC}^{Hybrid}[\psi(\boldsymbol{r})] = E_X^{local}[n(\boldsymbol{r})] + \alpha\left(E_X^{HF}[\psi(\boldsymbol{r})] - E_X^{local}[n(\boldsymbol{r})]\right) + E_C^{local}[n(\boldsymbol{r})] \quad (2.38)$$

Examples of hybrid functional including PBE0 where α=0.25, the Heyd-Scuseria-Ernzerhof (HSE) where the functional is revised as [11, 12]

$$E_{XC}^{HSE} = \alpha E_X^{HF}(\omega) + (1-\alpha)E_X^{PBE,SR}(\omega) + E_X^{PBE,LR}(\omega) + E_C^{PBE}(\omega) \quad (2.39)$$

The Coulomb potential is decomposed into the long-range term and the short-range term by an error function with range separation parameter ω.

$$\frac{1}{r} = \frac{erfc(\omega r)}{r} + \frac{\mathrm{erf}(\omega r)}{r} \quad (2.40)$$

where erf is error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt \quad (2.42)$$

and erfc(x)=1-erf(x).

The popular HSE06 functional sets α=0.25, ω=0.2. When ω=0, HSE goes back to PBE0.

The screened exchange (sX) functional is non-local with a combination of a short-range fraction of HF exchange and a semi-local model of long-range exchange [13]. The screened exchange energy is

$$E_X^{sX}[\psi(\boldsymbol{r})] = -\frac{e^2}{8\pi\varepsilon_0}\int d\boldsymbol{r}d\boldsymbol{r}' \sum_{\boldsymbol{i},j,k,q} \frac{\psi_i^*(\boldsymbol{r})\psi_j^*(\boldsymbol{r}')\psi_i(\boldsymbol{r}')\psi_j(\boldsymbol{r})e^{-k_{TF}|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{2.43}$$

where i and j label electronic bands, k and q are the k points and k_{TF} is Thomas-Fermi screening. In reciprocal space, the potential is also decomposed into a short-range and long-range part.

$$\frac{4\pi e^2}{|\boldsymbol{G}|^2} = \frac{4\pi e^2}{|\boldsymbol{G}|^2 + k_{TF}^2} + \left(\frac{4\pi e^2}{|\boldsymbol{G}|^2} - \frac{4\pi e^2}{|\boldsymbol{G}|^2 + k_{TF}^2}\right) \tag{2.44}$$

For typical semiconductor, a Thomas-Fermi screening length is about 1.8Å^{-1} which comes from Debye length at 300K, dielectric constants and valence electron density. This value yields reasonable band gaps. For $k_{TF} \to \infty$, the exchange-correlation term goes to the free electron limit of LDA. If $k_{TF} = 0$, it goes to pure HF.

Only norm-conserving pseudopotentials can be used in Hybrid functional at present, therefore, a high energy cut-off is needed. The calculation time is much longer than LDA and GGA. Fig 2-5 shows the comparison of band gaps obtained from experiment, sX, and GGA. It is revealed that sX's result is much better than non-hybrid functional.

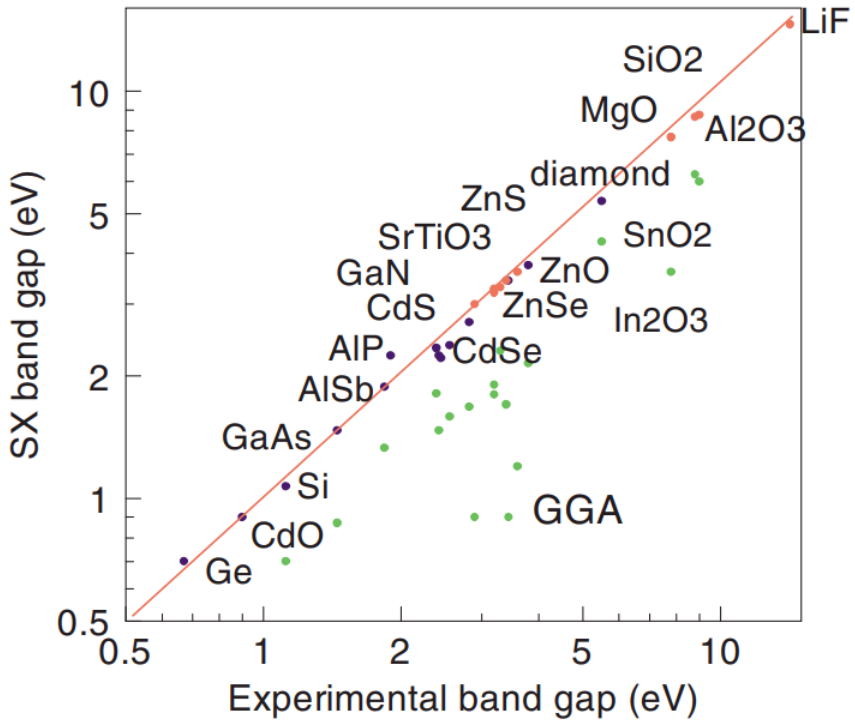


Fig 2-5. Comparison of the calculated sX and experimental minimum band gaps for various compounds are given. The GGA-PBE values are also given, from ref [12].

2.6.4 Geometry Optimization and band structure calculation

In section 2.2, we explained how to conduct electron minimization. After each step of SCF, we need to judge whether the current structure is the most stable structure. The force and stress are necessary to be calculated. According to Hellman-Feynman theorem [14, 15], if we know the electron density we can derive all the forces in system.

$$\boldsymbol{F} = -\nabla_{\boldsymbol{R}} E = -\nabla_{\boldsymbol{R}}[< \psi|\boldsymbol{H}|\psi >] = -< \psi|\nabla_{\boldsymbol{R}}\boldsymbol{H}|\psi > \tag{2.45}$$

We can obtained the stress tensor by

$$\sigma_{\alpha\beta} = \frac{1}{V}\frac{\partial E}{\partial \varepsilon_{\alpha\beta}} \tag{2.46}$$

Where V is the volume of the unit cell. The relaxation of crystals is a quasi-Newton method known as BFGS method [16]. As the force and tensor (if you relax the cell size as well) converge, the geometry optimization is finished.

Then we can obtain the electron density, the total energy and eigenstates and eigenvalues (energy level) of the final structure. The band structure is calculated by plotting the energy level in each k point sampled in the first Brillouin zone.

2.7 Corrections

2.7.1 Van der Waals dispersion correction

The van der Waals interaction is long-range dipole-dipole interaction which is weaker than covalent bond and hydrogen bond but very important as it is the force between the layered materials, namely van der Waals heterostructures. The weakness of the force makes it possible to peel a monolayer off the bulk materials like TMD or BN by mechanical exfoliation, without leaving dangling bonds in the monolayer. Besides, the van der Waals force is an important component of intermolecular binding energies and surface adsorption. However, DFT does not take van der Waals or hydrogen bond into account, so a semi-empirical method is needed to correct the total energy, the surface landscape, as well as the wave function.

The most popular correction schemes are by Grimme [17] and Tkatchenko-Scheffler (TS) [18]. Grimme provides highly empirical parameters covering most of the elements, while TS reduces the number empiric parameter to one in the damping function, with asymptotic accuracy to 5%. The correction energy functional of TS is given as [19]

$$E_{dis} = -\frac{1}{2}\sum_{j>i}\frac{C_{6ij}}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^6}\frac{1}{1+\exp\left(-d\left(\frac{|\boldsymbol{r}_i - \boldsymbol{r}_j|}{s_R R_0} - 1\right)\right)} \tag{2.47}$$

C_{6ij} is the dispersion coefficient for two atoms i and j. There are 5 parameters need to set up in our calculation. C_6, R_0, α are different for each element and s_R, d can be determined

universally, which by default is set to s_R=0.94, d=20. We note that a damped function is necessary for the energy correction functional so that if the distance between two atoms i and j is too small, the functional is chosen to decay to zero. This is to assure that below the van der Waals radii R_0, the dispersion correction is negligible so that normal bonds are not affected.

2.7.2 Lany-Zunger Scheme of calculating charged point defects

In chapter 1 we introduced basic concept and different kinds of point defects. Aside from being a local charge trap and scattering centre, defects can also cause Fermi level pinning. It is extremely important to study the formation of a defect and determine how easily a defect can form in a pristine structure. The way to calculate how a defect form is to calculate its formation energy. This is defined as the difference of total energies between defect-containing structure and the defect-free structure considering the chemical potentials of defect species. Therefore, the formation energy is

$$H_q = E_q - E_H + \sum_{\alpha} n_\alpha \mu_\alpha \tag{2.48}$$

where E_q is the energy of defect system with charge q, E_H is the energy of the defect-free structure, the third term is the chemical potential of each defect species.

The definition is straightforward but to calculate numerically is not as easy as it seems to be. DFT can only calculate the total energy of the periodic structure, however, a defect breaks the translation symmetry so if you create a supercell with a point defect, calculate the total energy of the supercell, the total energy is not the energy of the system with an isolated defect but the energy of a periodic distributed point defect system. If the defect has charge, then the Coulomb interaction between the defect and its periodic images is taken into account as you calculate the total energy. Therefore, we need to find a way to cancel out the errors induced by periodic charge defects.

Lany and Zunger [20] designed a correction method shown in Fig 2-6. We can split the term E_q-E_H into two parts by adding a term of the total energy of defect-free structure with charge q. Although the absolute value of total energy calculated by DFT is meaningless, the

energy difference is meaningful and convergent. In the first column, the energy difference cancels out the effect of mirror charges interaction. Therefore, all you need is to calculate the total energy of charged periodic defect supercell and subtract the total energy of supercell and charge without defect from it. The second column is the energy difference of a charged pristine system and neutral pristine system, which may induce the mirror-image problem. However, this value is just the energy of moving an electric charge from far away to the defect centre in the supercell. At the infinite distance, the charge's energy level is at Fermi level, whereas in defect centre the energy level should be at valence band maximum (VBM) if it is positive, at conduction band minimum (CBM) if it is negative. Therefore, the formation energy should be

$$H_q = E_q - E_m + q\big(E_F - E_C U(-q) - E_V U(q)\big) + \sum_\alpha n_\alpha \mu_\alpha \tag{2.49}$$

where E_C is CBM, E_V is VBM, E_F is Fermi level, U(q) is the Heaviside function which is 1 when q>0 and 0 when q<0.

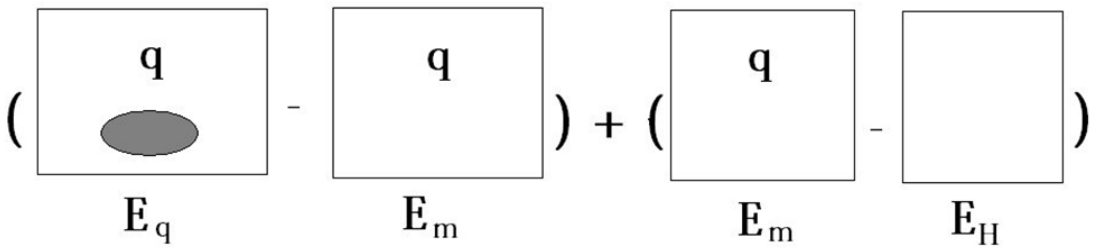


Fig 2-6. Illustration of Lany-Zunger scheme: by adding an intermediate reference term, the energy difference can split into two convergent and DFT accessible parts.

Firstly, it is noteworthy that Lany-Zunger scheme can correct errors induced by Coulomb interaction between periodically charged defects and it is independent of the size of the supercell. Even though, you still need to create the supercell large enough so that the defect cannot interact with its periodic image and the system can be treated as with an isolated defect. You should do this even in a neutral defect system, which is the same spirit as the need to create a vacuum with enough length if you want to simulate a surface. In most cases,

the defect density is low enough to be treated as mutually isolated to each other. For example, the density of S vacancies in 2D MoS_2 is ~$10^{13}cm^{-3}$ [21, 22], which is 10 defect per μm^{-3}. Therefore, we need to construct a supercell about 10000Å to contain 10 S vacancies, which is not realistic for DFT calculation. Nevertheless, we know that for S vacancy in MoS_2 we can treat each S vacancy as isolated to others, so we need to create an large enough supercell of MoS_2 for defect calculations. There are exceptions in the highly doped compound or alloys, the doping species can be treated as an ingredient rather than defects.

Secondly, as the size of supercell becomes larger and approaches to infinity, the formation energy calculated becomes more accurate. If you are not sure about how to choose the size of a supercell, a standard protocol is to calculate the formation energy from small size to large and observe the convergence of formation energy.

2.7.3 Conclusion

In this chapter, we introduce the theoretical methods used in this work, the DFT. In this work, most of the projects are completed by using Cambridge Serial Total Energy Package (CASTEP) code to perform DFT calculation. CASTEP was firstly written in Fortran 77 by M. C. Payne and improved by M. D. Segall et.al and rewritten in Fortran 90. Details of the code are summarized in ref [23].

Reference

1. M Born, J R Oppenheimer, Annalen der Physik **389** 457 (1927).
2. J A Pople, et al. International Journal of Quantum Chemistry **14** 545 (1978).
3. L Pietronero and F Zirilli, Nuov Cim B **24** 57 (2007).
4. W Kohn, L J Sham, Phys. Rev. **140** A1133 (1965).
5. M C Payne, et al. Rev. Mod. Phys. **64** 1045 (1992).
6. C Kittel, Introduction to Solid State Physics. New York: Wiley (1996).
7. H J Monkhorst and J D Pack, Phys. Rev. B **13** 5188 (1976).
8. D Vanderbilt, Phys. Rev. B **41** 7892 (1990).
9. J P Perdew, et al. Phys. Rev. B **46** 6671 (1992).
10. M Cococcioni, The LDA+ U approach: a simple Hubbard correction for correlated ground states[J]. Correlated Electrons: From Models to Materials Modeling and Simulation, 2012, 2: 4.4-4.40.
11. K Hummer, J Harl, and G Kresse, Phys. Rev. B **80** 115205 (2009).
12. J Heyd, G E Scuseria, and M Ernzerhof, J Chem. Phys **118** 8207 (2003).
13. S J Clark and J Robertson, Phys. Rev. B **82** 085208 (2010).
14. H Hellmann, Leipzig: Franz Deuticke 285 (1937).
15. R P Feynman, Phys. Rev. **56** 340 (1939).
16. B G Pfrommer, M Cote, S G Louie and M L Cohen, J. Comp. Phys. **131** 233 (1997).
17. S Grimme, J Comp. Chem. **25** 1463 (2004).
18. A Tkatchenko and M Scheffler, Phys. Rev. Lett **102** 073005 (2009).
19. T Bučko et al. Phys. Rev. B **87** 064110 (2013).
20. S Lany and A Zunger, Phys. Rev. B **78** 235104 (2008).
21. W Zhou, X Zou, S Najmaei, Z Liu, Y Shi, J Kong, J Lou, P M Ajayan, B I Yakobson, and J C Idrobo, Nano Lett. **13** 2615 (2013).
22. J Hong, Z Hu, M Probert, K Li, D Lv, X Yang, L Gu, N Mao, Q Feng, L Xie, J Zhang, D Wu, Z Zhang, C Jin, W Ji, X Zhang, J Yuan and Z Zhang, Nat. Comm. **6**, 6293 (2015).
23. M D Segall, et al. Journal of Physics: Cond. Matt. **14** 2717 (2002).

Chapter 3 Chemical trends of Schottky barrier behaviour on monolayer hexagonal B, Al and Ga nitrides

The Schottky Barrier Heights (SBH) of metal layers on top of monolayer hexagonal X-nitrides (X=B, Al, Ga, h-XN) are calculated using supercells and density functional theory to understand the chemical trends of contact formation on graphene and the 2D layered semiconductors such as the transition metal dichalcogenides. The Fermi level pinning factor S of SBHs on h-BN is calculated to be near 1, indicating no pinning. For h-AlN and h-GaN, the calculated pinning factor is about 0.63, less than for h-BN. We attribute this to the formation of stronger, chemisorptive bonds between the nitrides and the contact metal layer. Generally, the h-BN layer remains in a planar sp^2 geometry and has weak physisorption bonds to the metals, whereas h-AlN and h-GaN buckle out of their planar geometry which enables them to form the chemisorptive bonds to the metals.

3.1 Background

Two-dimensional (2D) layered semiconductors have been proposed for use in future electronic devices including tunnel field effect transistors (FETs) and low power sensors [1-3]. However, their device performances tend to be limited by contact resistances, due to the presence of Schottky barriers at their contacts [4, 5]. In some cases such as 2D MoS_2, the Schottky barrier heights (SBHs) show strong Fermi level pinning (FLP) [5-11], which limits our ability to choose a contact metal with a suitable work function so as to reduce the SBH. In other cases like h-BN, there is little FLP according to theoretical calculations [12]. The behaviour of SBHs in two-dimension systems is often attributed to their dimensionality and to the presence of van der Waals interlayer bonding [13-15]. However, it is interesting to

understand how much of the behaviour depends on dimensionality and how much depends on the chemical bonding between the layers. We therefore study the chemical trends in the system of the 2D hexagonal group III-nitrides, h-BN, h-AlN and h-GaN, which show both types of behaviour.

In general, FLP can arise from either intrinsic or extrinsic effects [9, 16]. The extrinsic effects are due to point defects such as vacancies created by the formation of the contacts, which in principle can be avoided and will not be considered further here. The intrinsic mechanism is due to charge transfer between the metal and the nitride, via the traveling wave states of the metal extending into the semiconductor band gap, where they are called 'metal-induced gap states' or MIGS [16-19]. The MIGS can pin the Fermi level if their density is large enough and if they have not decayed too much across any bonding 'gap' between the contacts and the semiconductor layer [6, 13]. Thus, the question turns out to be: what is the actual bonding between the contact metal and 2D layer in each case, and how does it vary in the nitrides?

3.2 Methods

The calculations of contacts on 2D materials are carried out using the CASTEP density functional theory (DFT) code [20], using plane waves, pseudopotentials and model periodic supercell models of the systems. We use the Perdew-Burke-Ernzerhof (PBE) form of generalized gradient approximation (GGA) for the exchange-correlation functional. Norm-conserving pseudopotentials were used for this purpose with a plane wave cut-off of 490eV. The screened exchange (SX) hybrid functional is used to overcome the band gap errors of the GGA functional [21]. A correction to the GGA treatment of the van der Waals interaction is included by using the Tkatchenko-Scheffler (TS) scheme [22, 23] in order to obtain the better bond lengths.

3.2.1 Lattice match

In each case, we created a supercell with a close degree of lattice matching between the nitride layer and the metal contact layers. We used six layers of metal in a hexagonal cell (111

or 0001). The vertical size of the supercell includes a vacuum gap of z=30Å (for X=B) or 40Å (for X=Al, Ga). We fix the cell constant in XN and strain the cell of metals, which is consistent with the experiment where metals are deposited onto the nitride layers. All possible hexagonal supercells out of the primitive one can be created and the length of them are:

$$a_{mn} = a_0\sqrt{m^2 + n^2 + mn} \qquad (m, n = 0, 1, 2, 3 \ldots) \tag{3.1}$$

where a_0 is the length of primitive cell and a_{mn} gives the supercell length from small to large. The lateral size of the supercell is chosen to attain a reasonable degree of lattice matching between the nitride and the metal layers, while not being so large that it leads to excessive computational costs. Our supercells provide a lattice matching to within 5%. Table 3-1 lists the lattice vectors of the metal lattice and the nitride lattice to achieve this matching. The lattice mismatch for each case is given in Tables 3-2, 3-3 and 3-4.

Table 3-1. In-plane matching of metal and h-XN lattices in each supercell. For example, the h-BN/Sc contact $2 \times 2 = \sqrt{7} \times \sqrt{7}$ means 2×2 double size supercell of hexagonal Sc (0001) surface is fitted into a $\sqrt{7} \times \sqrt{7}$ supercell of h-BN.

Transition metal	*h-BN*	*h-AlN*	*h-GaN*
Sc(0001)	$2 \times 2 = \sqrt{7} \times \sqrt{7}$	$1 \times 1 = 1 \times 1$	$1 \times 1 = 1 \times 1$
Ag(111)	$\sqrt{3} \times \sqrt{3} = 2 \times 2$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$
Al(111)	$\sqrt{3} \times \sqrt{3} = 2 \times 2$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$
Ti(0001)	$\sqrt{3} \times \sqrt{3} = 2 \times 2$	-	-
Cu(111)	$1 \times 1 = 1 \times 1$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$	$\sqrt{7} \times \sqrt{7} = 2 \times 2$
Co(0001)	$1 \times 1 = 1 \times 1$	$\sqrt{7} \times \sqrt{7} = 2 \times 2$	$\sqrt{7} \times \sqrt{7} = 2 \times 2$
Pd(111)	$\sqrt{3} \times \sqrt{3} = 2 \times 2$	-	-
Ni(111)	$1 \times 1 = 1 \times 1$	$\sqrt{7} \times \sqrt{7} = 2 \times 2$	$\sqrt{7} \times \sqrt{7} = 2 \times 2$
Pt(111)	$\sqrt{7} \times \sqrt{7} = 3 \times 3$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$	$2 \times 2 = \sqrt{3} \times \sqrt{3}$
MoO_3	$4 \times 6 = 3 \times 5\sqrt{3}$	$4 \times 3 = 5 \times 2\sqrt{3}$	$5 \times 3 = 6 \times 2\sqrt{3}$

Table 3-2. Details of h-BN/metal contact. The lattice mismatch is the difference between the lengths of two supercells divided by length of the bigger supercell. The binding energy is the energy per formula unit for binding the two materials together. The equilibrium distance is the normal plane - the top plane separation between the h-BN layer and the surface of metals.

Transition metal	*Work function (eV)*	*Lattice mismatch (%)*	*Binding energy (eV/ XN)*	*Layer distance (Å)*	*P-type SBH(eV)*
Sc(0001)	3.50	1.17	-0.19	3.57	*3.43*
Ag(111)	4.26	1.11	-0.85	3.05	*2.74*
Al(111)	4.28	1.99	-0.22	3.22	*2.59*
Ti(0001)	4.33	0.99	-0.53	2.01	*2.46*
Cu(111)	4.65	1.01	-0.19	2.99	*2.13*
Co(0001)	5.00	0.91	-0.39	3.05	*1.78*
Pd(111)	5.12	5.83	-0.24	2.36	*1.83*
Ni(111)	5.15	1.51	-0.27	3.02	*1.71*
Pt(111)	5.65	3.29	-0.01	3.51	*1.51*
MoO_3	6.61	4.23(x) 1.20(y)	-0.22	2.85	*-0.03*

Table 3-3. List of details of the h-AlN/metal contact.

Transition metal	*Work function (eV)*	*Lattice mismatch (%)*	*Binding energy (eV/XN)*	*Layer distance (A)*	*P-type SBH(eV)*
Sc(0001)	3.50	5.3	-1.22	2.30	*2.15*
Ag(111)	4.26	4.4	-0.89	2.70	*1.68*
Al(111)	4.28	5.3	-0.73	2.04	*2.30*
Cu(111)	4.65	5.7	-0.74	2.22	*1.65*
Co(0001)	5.00	5.6	-1.14	2.05	*1.23*
Ni(111)	5.15	5.0	-1.14	2.04	*1.32*
Pt(111)	5.65	2.3	-1.02	2.24	*1.00*
MoO_3	6.61	1.23(x) 2.05(y)	-0.35	1.90	*0.28*

Table 3-4. List of details of the h-GaN/metal contact.

Transition metal	*Work function (eV)*	*Lattice mismatch (%)*	*Binding energy (eV/XN)*	*Layer distance (A)*	*P-type SBH(eV)*
Sc(0001)	3.50	2.8	-1.34	2.21	*1.87*
Ag(111)	4.26	3.6	-0.64	2.67	*1.65*
Al(111)	4.28	2.7	-0.67	1.99	*1.84*
Cu(111)	4.65	4.9	-0.70	2.18	*1.31*
Co(0001)	5.00	3.0	-1.19	2.04	*0.99*
Ni(111)	5.15	2.4	-1.63	2.03	*1.35*
Pt(111)	5.65	0.4	-0.95	2.17	*0.67*
MoO_3	6.61	2.60(x) 0.48(y)	-0.31	2.20	*0.04*

3.2.2 Core levels

We calculated the relaxed atomic structure and then calculated the p-type Schottky barrier height (SBH), which is the energy difference between metal Fermi level and valence band maximum (VBM) of the semiconductor. It can sometimes be difficult to identify the semiconductor VBM energy in this type of calculation because of the interaction between the semiconductor valence states with the metal states. Sometimes, the semiconductor's residual band structure can be used as a reference marker, as, for example, in Gong [7]. In other cases, the vacuum level can be used as a reference level to identify a dipole, as in Bokdam [12]. Here, we use a core level of the semiconductor as a reference energy to identify the nitride VBM energy in the interacting system (Fig. 3-1). This is the analog of Kraut's method in photoemission spectroscopy [24]. This method was previously used by us to calculate the

Schottky barrier heights for transition metal dichalcogenides (TMD). [11] The $1s^2$ state of boron, $2s^2$ of aluminium and $3s^2$ of gallium are included in the valence shell for this purpose. Ultra-soft pseudopotentials were generated for this purpose. The plane wave cut-off is tested to be 250eV, 280eV and 450eV for B, Al, Ga respectively. For each system, the cut-off energy is chosen to be that necessary to converge the total energy to less than $5x10^{-6}$ eV/atom. Fig. 3-2 shows the averaged electrostatic potential for a supercell, which allows us to define the vacuum energy.

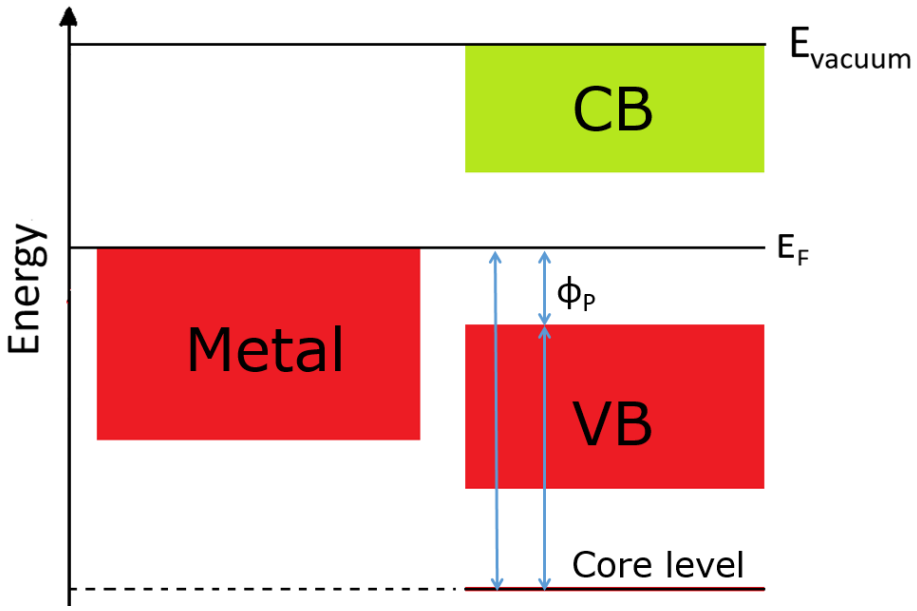


Fig. 3-1 Schematic of core level reference method. Energy bands are shown as colored blocks.

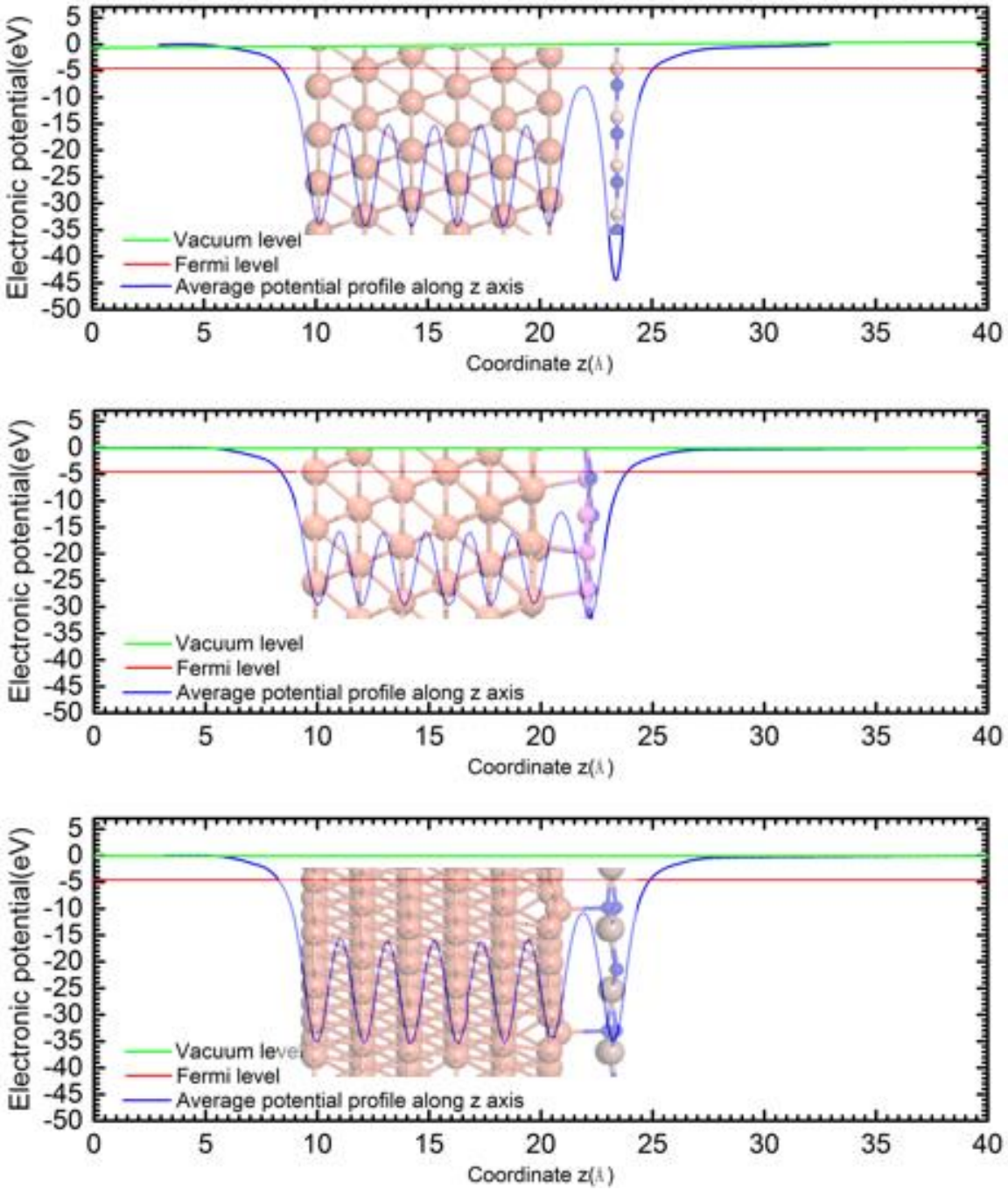


Fig 3-2. The averaged electronic potential of h-BN, h-AlN and h-GaN monolayers adsorbed on six layers of Cu (111) surface. The green line is vacuum level and the red line is the Fermi level. The potential level coincident with vacuum level in vacuum region, which shows the vacuum is thick enough. Inside the layers, the potential oscillates with the atom layers. Note the longer decay length metal-induced gap states between the metal and h-BN layer than for other nitrides.

3.3 Results

3.3.1 Structure and bands of 2D nitrides

The equilibrium structures of isolated monolayer h-XN are relaxed and the calculated lattice parameters are 2.53Å (h-BN, 1.19% from experimental result [25]), 3.13Å (h-AlN, 0.20% errors [26]) and 3.22Å (X=Ga).

There are various possible high symmetry bonding geometries between the metal (Me) layers and the nitride that maximize the interaction between these layers, as shown in Fig 3-3. Fig 3-3(a) shows three possible configurations for large metal atoms like Sc; the Metal/N C_{3v} symmetry on-top site, the C_s hollow site, and the C_s bridge site. For metals like Ag, there are two main sites, Fig 3(b); the C_s bridge site, and the C_s on-top/hollow site combination. Other cases have lower symmetry. The energetically most favorable binding configurations are calculated.

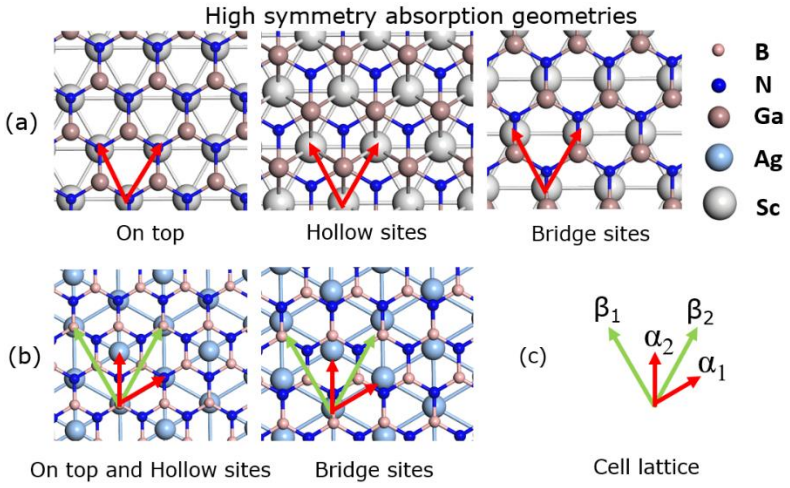


Fig 3-3. Top view of possible top contact binding configurations with symmetries. Only the top layer of metal s are shown. (a) is the 1×1 h-GaN matching with 1×1 Sc (0001) surface, (b) is the 2×2 h-BN matching with $\sqrt{3} \times \sqrt{3}$ Ag (111) surface. The color of the atom, B=peach, N=blue, Sc=light gray, Ga=brown, Ag=light blue. (c) The 1×1 cell lattice α of metal and GaN in (a) and metal in (b) is marked red, while the 2×2 BN lattice or the $\sqrt{3} \times \sqrt{3}$ Ag lattice β in (b) is marked green.

The binding energy per molecular unit of different configurations is calculated and are given in Tables 3-2 to 3-4. Due to the high electronegativity of N and its preference to bond to the contact's metal atoms, the most favorable position for N and X are the on-top and hollow sites, respectively. Therefore, in the 1×1 metal supercells in Fig. 3-3(a), the N on-top site is favoured while for the $\sqrt{3} \times \sqrt{3}$ metal supercells in Fig. 3-3(b), the on-top and hollow sites are favoured. It is always possible to make as many N atoms on top of metal atoms while keeping the X away from the metal for other cases in Table 3-1.

The differences of binding energies between different symmetry configurations are also compared with the thermal perturbation energy. For example, in the Ni|h-AlN and Ni|h-GaN contacts, the planar structures deform most, the N-Ni bonds are shortest and strongest, the largest difference of binding energies between N on top and other configurations are 10.2meV and 8.2meV per formula unit respectively. The thermal ripple kT in room temperature is ~25.6meV per formula unit, which is larger. Thus it is concluded that the probability of finding each symmetry configuration in experiment is nearly the same for all nitrides.

Fig. 3-4 shows a side view of the layer contact of h-XN with various metals. We show a few cases with the shortest and longest equilibrium distances as examples. The binding energies and layer distances of the most favorable binding sites are given in Tables 3-2 to 3-4.

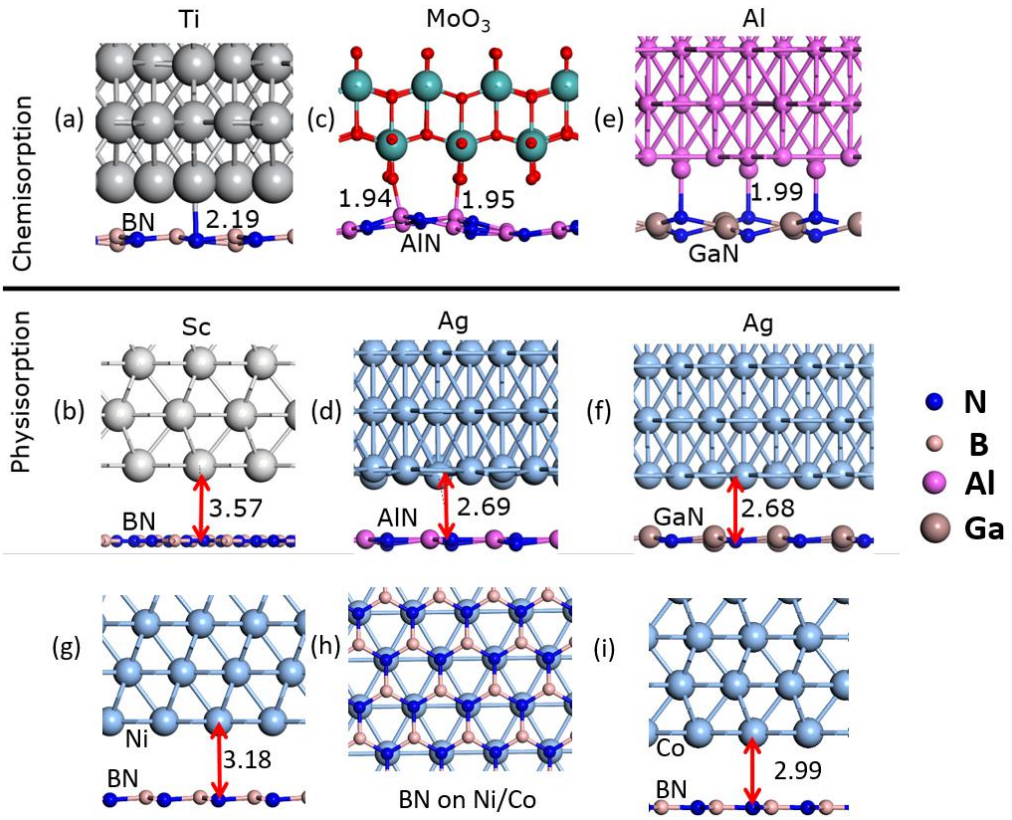


Fig 3-4. Top contact bonding at (a)/(b) Ti/Sc-BN interface, (c)/(d) MoO_3/Ag-AlN interface and (e)/(f) Al/Ag-GaN interface. The color of atom, B=peach, N=blue, O=red, Al=fuchsia pink, Sc=light grey, Ti=dark gray, Mo=cyan, Ga=brown, Ag=light blue. The shortest distances are marked in black dot line. The upper half is classified as chemisorption while the half beneath is physisorption, with no bond.

3.3.2 Schottky Barrier Height and Fermi level pinning

We now consider the calculated values of Schottky barrier height using PBE as the exchange-correlation functional, as given in Tables 3-2 to 3-4. These are plotted against experimental metal work functions of the metals in Fig 3-5. In order to display three semiconductors SBH on the same diagram, we align their band edges according to their charge neutrality levels (CNLs). The CNL is defined as the branch point energy where the Greens function of the density of states is 0 [16, 27]. Table 3-5 lists the band gaps and CNLs for the hexagonal XN phases. The resulting band gaps in sX are 5.76eV (h-BN), 3.97eV (h-AlN) and

2.02eV (h-GaN). Note that these band gaps are smaller than for the sp^3 phases. The CNLs lie in the lower half of the gap.

Table 3-5. Comparison of band gaps calculated by PBE and sX-LDA and CNLs extracted from sX bands

Semiconductor	**Band gap (eV)**		**Isotropic CNL(eV), sX**
	PBE	sX-LDA	
ML h-BN	4.77	5.72	2.12
ML h-AlN	2.92	3.84	2.18
ML h-GaN	1.88	3.19	1.33

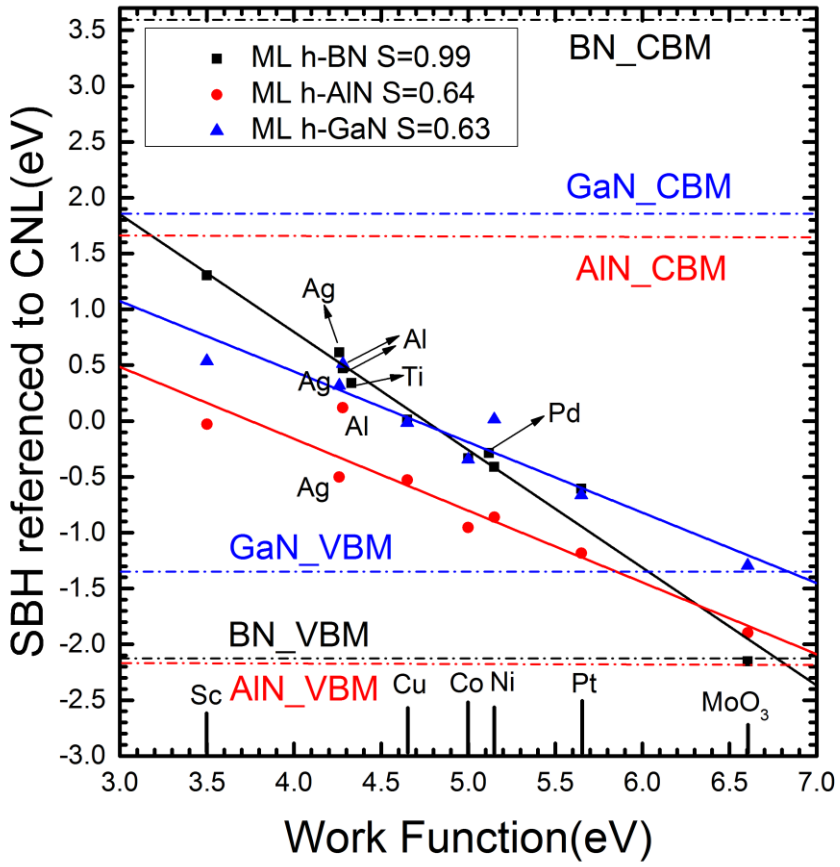


Fig 3-5. P-type SBH of metal-h-XN contact, referred to the CNLs of each XN's CNLs, colors of data points and their band edges, with an ideal linear fitting with pinning factor of 0.99 (X=B), 0.64 (X=Al), 0.63 (X=Ga) and standard error of 0.06, 0.09, 0.08, respectively.

The transition metals used for contacts are the Sc, Co, Ti (0001) surfaces, the Ag, Al, Ni, Pd, Pt, Cu (111) surfaces and MoO_3. The experimental work function values are taken from Michaelson [29]. The work functions range from 3.5eV for Sc to 6.6 eV for MoO_3 [28, 29]. Despite the different metal reactivities, to be discussed below, the p-type SBH shows a surprisingly good linear relationship against metal work function. The absolute value of slope represents the pinning factor S. Within the MIGS model of pinning, the n-type SBH is given by [18],

$$\phi_n = S(\Phi_M - \mathrm{E}_{cnl}) + E_{cnl} - \chi_s \qquad (3.2)$$

where χ_s is the electron affinity of the semiconductor, Φ_M is the work function of the metal contact, and E_{cnl} is the CNL of the semiconductor. The pinning factor S can vary from 0 (Bardeen limit) to 1 (Schottky limit). S=1 means no pinning in the interface and S close to 0 means strong pinning. S can be interpreted in terms of the density of interface gap states which cause pinning. A large density of gap states will result in a smaller S and more pinning.

The calculated pinning factor of ML h-BN is 0.99, while that of ML h-AlN is 0.64 and ML h-GaN is 0.63. This shows that the Fermi level pinning does not occur in h-BN while it is significant in h-AlN and h-GaN. The result of S=0.99 for h-BN is consistent with Bokdam et al [12]. However, the dependence of barrier height vs metal work function found here is much more linear and covers a wider range of work functions.

3.3.3 Chemical trend

We now explain how FLP arises in these systems. Pinning requires some charge transfer from the metal to the nitride layer to create a dipole that opposes the change in the work function of the metal. The charge transfer occurs through the overlap of states of the metal and the nitride layer. These states are generally the MIGS. If the separation between the nitride and metal layers is too large, the MIGS will have decayed too much at the nitride layer to allow much charge transfer, and the pining will be small, with S ~1, Fig 3-4. On the other hand, if the layer separation is smaller, there can be stronger coupling by chemisorptive bonds between the nitride layer and the metal, more charge transfer, and this would give stronger pinning.

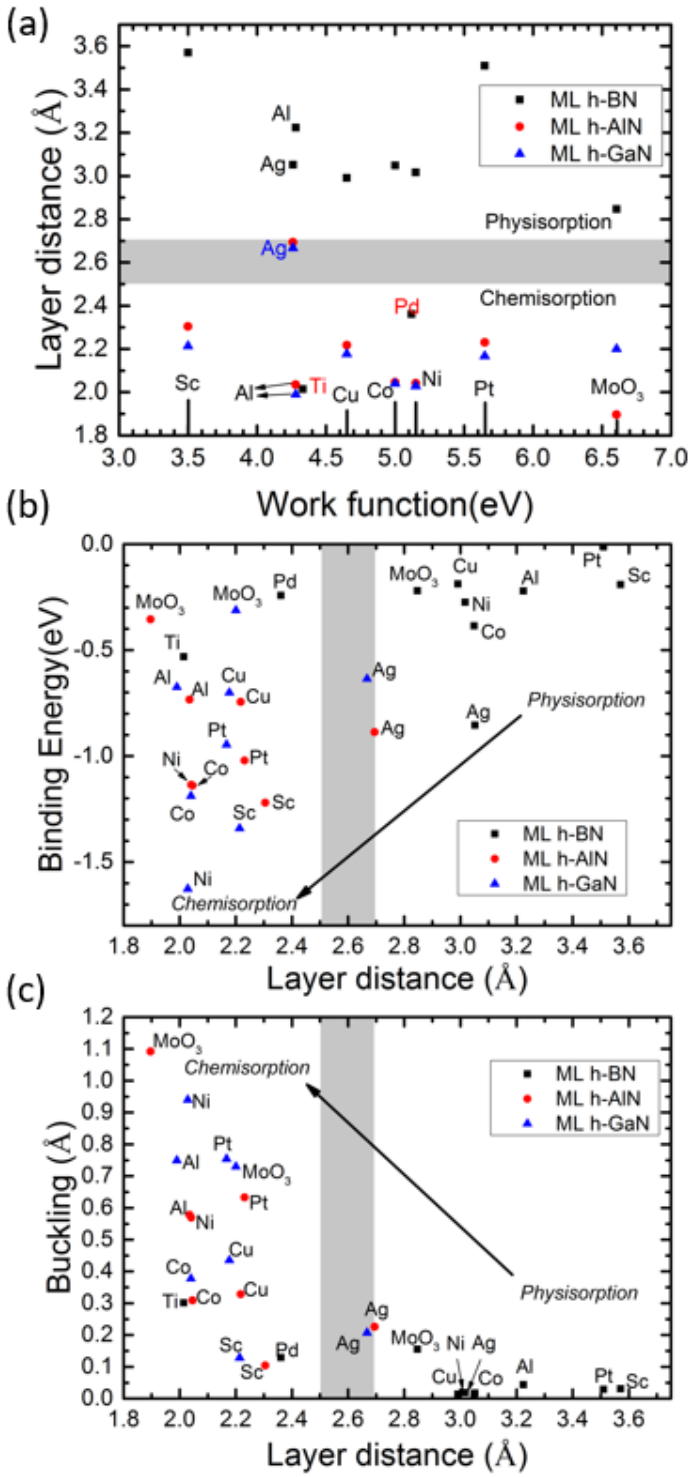


Fig. 3-6 (a) Metal work function (W) vs equilibrium distances between XN and metal layers (D). (b) D vs buckling. (c) D vs binding energies. Metal and N form bond when XN is chemisorbed onto metals, except the case of MoO_3 O-rich, where metal

and oxygen tend to form bonds. The grey region shows a transition from physisorption to chemisorption. For X=B, the distances are mainly above the grey region. Ti and Pd are two exceptions, marked as red in (a). For X=Al and Ga, the distances are mainly below the grey region. Exceptions are Ag which is marked as blue in (a).

Fig 3-6(a) plots the metal to nitride layer separation against the work function of the metal for all three nitrides. Despite the preference of metal atoms to bond with the nitrogen sites, interestingly there is not much dependence of the separation on the metal work function. Generally, the data points for h-BN stand out from those for h-AlN and h-GaN.

Fig 3-6(b) plots the binding energy of nitrides to the metal layers against the interlayer separation, d. A strong dependence is seen. There is weak binding for d > 2.7Å, but increasing strong binding for d< 2.5Å. Clearly, 2.6Å separates the weakly bound physisorbed layers from the strongly bound chemisorbed layers.

But why do most metals on h-BN fall into the physisorbed category? Fig 3-6(c) plots the binding energy against the buckling distance of the nitride layer when there is metal on top of it. It is clear that there is only strong binding of the metal layer and short interlayer separation if there is buckling of the nitride layer. The fundamental reason is that h-BN is more stable in its planar sp^2 bonded state, with no buckling. In contrast, h-AlN and h-GaN, like their bulk phases, are more stable in their sp^3 bonded phases, consistent with buckling. The formation of a short, strong bond between the metal and a nitrogen site, making the N 4-fold coordinated, requires the buckling.

We see that there are a few exceptional cases. Ti and Pd are more reactive metals. These are able to buckle h-BN. On the other hand, despite its rather low work function, Ag is a noble metal and it is rather unreactive. It is in the weakly bonded category and causes less buckling on h-AlN or h-GaN than other metals of this work function. A planar sp^2 bonded h-XN layer structure is preserved and there is no direct bond formation. Apart from the metal-O bonds in the case of O-rich MoO_3, metal-N bonds dominate in Fig. 3-4.

MoO_3 consists of a bilayer of MoO sites with 3-fold and 2-fold bonded O sites, 6-fold bonded Mo sites, plus two external layers of monovalent O sites on each side. [24] The interfaces of MoO_3 to the nitrides involve contact metal bonds to the external layer oxygen sites of the MoO_3 which makes them 2-fold coordinated.

Finally, we can separate the interactions of metals with graphene, nitrides and the TMDs into weak, medium and strongly absorbing, as in Table 3-6. This uses the data of Popov [3] and Giovannetti [14] for graphene, and Kang *et al.* [8] and Guo *et al.* [11] for MoS_2. We see that the type of interaction is quite similar in each case.

Table 3-6. Three classes of adsorption types.

Semiconductor monolayer **Contact situation**	**Graphene[21]**	**TMD[25]**		**h-XN**		
	PBE	MoS2	WSe2	BN	AlN	GaN
Weak	Al, Cu, Ag, Au, Pt	In, Au	In, Au	Sc, Cu, Ag, Al, Co, Pt, Ni, MoO_3		
Medium	---	Pd	Ti, Pd		Ag	Ag
Strong	Co, Ni, Pd	Ti, Mo, W	Mo, W	Ti, Pd	Sc, Cu, Pt, Al, Co, Ni, MoO_3	Sc, Cu, Pt, Al, Co, Ni, MoO_3

Generally, the Fermi level pinning on the TMDs is stronger than on the nitrides for two reasons. First, the band gaps of the TMDs are smaller, so that the MIGS density of states is larger. Second, metal bonding to sulphur atoms of MoS_2 occurs directly, and it does not require any buckling. Sulphur is able to form extra bonds (over-coordinate) being a second-row element, whereas nitrogen is a first-row element which has less tendency to over-coordinate.

These three different behaviours occur all in nominally van der Waals bonded 2D systems. Thus, it is not enough to classify them as 2D or van der Waals solids, and it is necessary to consider in greater depth their bonding behaviours.

The bonding configuration of the h-BN configuration of h-BN on Cu(111) and Ni(111) with the metal above the nitrogen site by only a small energy difference is of interest for the chemical vapor deposition of h-BN on Cu [31] and also for spin-filtering using h-BN [32].

3.4 Conclusion

In summary, the p-type SBHs of 2D h-XN are calculated using DFT. FLP is found to occur on defect free h-AlN and h-GaN. The calculated p-type SBH values are consistent with MIGS theory, obeying a linear relationship with the metal work function. Where this occurs, the pinning can be attributed to chemisorptive bonding on the metal atoms and the N sites of the nitrides, which causes a buckling of the planar layers. The Fermi level in h-AlN and h-GaN is pinned near CBM, especially for h-GaN. The pinning factor for h-AlN and h-GaN are 0.64 and 0.63 respectively, lower than for h-BN. For h-BN, there is usually no pinning because of the stable B-N bond opposed buckling, while for h-AlN and h-GaN FLP is greater due to weaker in-plane bond. It is also found that the N-metal bond as well as equilibrium distance is relevant to FLP. The system with stronger interlayer chemical bond often has stronger FLP.

REFERENCES

1. S. Das, A. Prakash, R. Salazar, and J. Appenzeller, ACS Nano **8** 1681 (2014).
2. D. Sarkar, X. J. Xie, W. Liu, W. Cao, J. H. Kang, Y. J. Gong, S. Kraemer, P. Ajayan, and K. Banerjee, Nature **526** 91 (2015).
3. Y. C. Lin et al., Nat. Commun. **6** 7311 (2015).
4. Y. Du, L. Yang, H. Liu, and P. D. Ye, APL Mater. **2** 092510 (2014).
5. S. Das, H. Y. Chen, A. V. Penumatcha, and J. Appenzeller, Nano Lett. **13** 100 (2013).
6. I. Popov, G. Seifert, and D. Tomanek, Phys. Rev. Lett. **108** 156802 (2012).
7. C. Gong, L. Colombo, R. M. Wallace, and K. J. Cho, Nano Lett. **14** 1714 (2014).
8. J. Kang, W. Liu, D. Sarkar, D. Jena, and K. Banerjee, Phys. Rev. X **4** 031005 (2014).
9. D. Liu, Y. Guo, L. Fang, and J. Robertson, App. Phys. Lett. **103** 183113 (2013).
10. Y. Guo, D. Liu, and J. Robertson, App. Phys. Lett. **106** 173106 (2015).
11. Y. Guo, D. Liu, and J. Robertson, ACS Appl. Mater. Interface **7** 25709 (2015).
12. M. Bokdam, G. Brocks, M. I. Katsnelson, and P. J. Kelly, Phys. Rev. B **90** 085415 (2014).
13. Y. Liu, P. Stradins, and S. H. Wei, Sci. Adv. **2** e1600069 (2016).
14. G. Giovannnetti, P. A. Khomyakov, V. M. Karpan, J. van der Brink, and P. J. Kelly, Phys. Rev. Lett. **101** 026803 (2008).
15. P. A. Khomyakov, G. Giovannetti, P. C. Rusu, G. Brocks, J. van der Brink, and P. J. Kelly, Phys. Rev. B **79** 195425 (2009).
16. J. Robertson, J. Vac. Sci. Technol. B **18** 1785 (2000).
17. J. Tersoff, Phys. Rev. Lett. **52** 465 (1984).
18. W. Mönch, Phys. Rev. Lett. **58** 1260 (1987).
19. W. Mönch, Surf. Sci. **300** 928 (1994).
20. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, Z. Kristallogr. **220** 567 (2005).
21. S. J. Clark and J. Robertson, Phys. Rev. B **82** 085208 (2010).
22. A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102** 073005 (2009).
23. E. R. McNellis, J. Meyer, and K. Reuter, Phys. Rev. B **80** 205414 (2009).
24. E. A. Kraut, R. W. Grant, and J. R. Waldrop, Phys. Rev. Lett. **44** 1620 (1980).
25. R. W. Lynch and H. G. Drickamer, J. Chem. Phys. **44** 181 (1966).
26. P. Tsipas, S. Kassavetis, D. Tsoutsou, E. Golias, S. A. Giamini, C. Grazianetti, D. Chappie, A. Molle, M. Fanciulli, and A. Dimoulas, Appl. Phys. Lett. **103** 251605 (2013).
27. J. Robertson and B. Falabretti, J. Appl. Phys. **100** 014111 (2006).
28. H. B. Michaelson, J. Appl. Phys. **48** 4729 (1977).
29. Y. Guo and J. Robertson, Appl. Phys. Lett. **105** 222110 (2014).
30. A. W. Cowley and S. M. Sze, J. Appl. Phys. **36** 3212 (1965).
31. P. R. Kidambi, R. Blume, J. Kling, J. B. Wagner, C. Baehtz, R. S. Weatherup, R. Schlogl, B. C. Bayer, and S. Hofmann, Chem. Mater. **26** 6380 (2014).
32. V. M. Karpan, P. A. Khomyakov, G. Giovannetti, A. A. Starikov, and P. J. Kelly, Phys. Rev. B **84** 153406 (2011).

Chapter 4 Fermi level De-pinning for Dopant-free Silicon Solar Cells

In this chapter, we move from 2D nitride to more widely applied Si. Over 90% of the photovoltaic (PV) market is Si, despite all the alternative technologies [1]. The traditional Si PV device uses a p-n junction and contacts. However, in the effort to lower costs, there has been a desire to avoid doping due to its associated high processing temperatures. It is possible to design solar cell without dopants, using what are known as carrier-selective asymmetric contacts as electrodes, whose ideas come from organic light emitting diode (OLED) technology. De-pinning of the Fermi level (E_F) is shown to be critical for the operation of selective asymmetric electrode photovoltaic cells. Without de-pinning, the two electrodes would have a similar Schottky Barrier Height (SBH) and the output voltage would be close to zero. In this work, the degree of Fermi level pinning (FLP) is calculated by density function calculations for the cases of 0, 1 and 2 layers of a representative oxide HfO_2 inserted between a metal layer and a Si substrate. Two layers is shown to be enough to depin E_F sufficiently to allow typical electrodes of MoO_3 and nanostructured ZnO/silicate to give a potential difference as large as the Si band gap. The problem of Fermi level pinning (FLP) in metal-Si contact is therefore resolved to give a completely different rational of PV design.

4.1 Background

In recent years, new materials have emerged in the photovoltaic market such as perovskite and organic molecular, but over 90% of the PV market is still Si. Nevertheless, the detailed designs evolve rapidly. The p-n or p-i-n homo-junction is one of the simplest semiconductor device structures, and it has been used over many years in photovoltaic cells [1]. However, in the effort to lower costs, there is now a desire to avoid doping and its associated high processing temperatures or poisonous pre-cursor gases. It is possible to design solar cells without dopants, using what are known as carrier-selective asymmetric electrodes as contacts

[2-7]. These are concepts borrowed from organic light emitting diode (OLED) technology [8]. They use a light absorbing layer between two electrodes of very different work function. The electrodes serve to separate the electrons and holes to create a photocurrent or to inject electrons or holes into organic semiconductor to emit photons, shown in Fig 4-1. Previously, OLEDs were considered to have many difficult materials problems caused by, for example, the difficulty of doping organic semiconductors. This meant that high and low work function electrode materials were needed to inject carriers directly into the bands, and these electrode materials could be unstable or atmosphere sensitive.

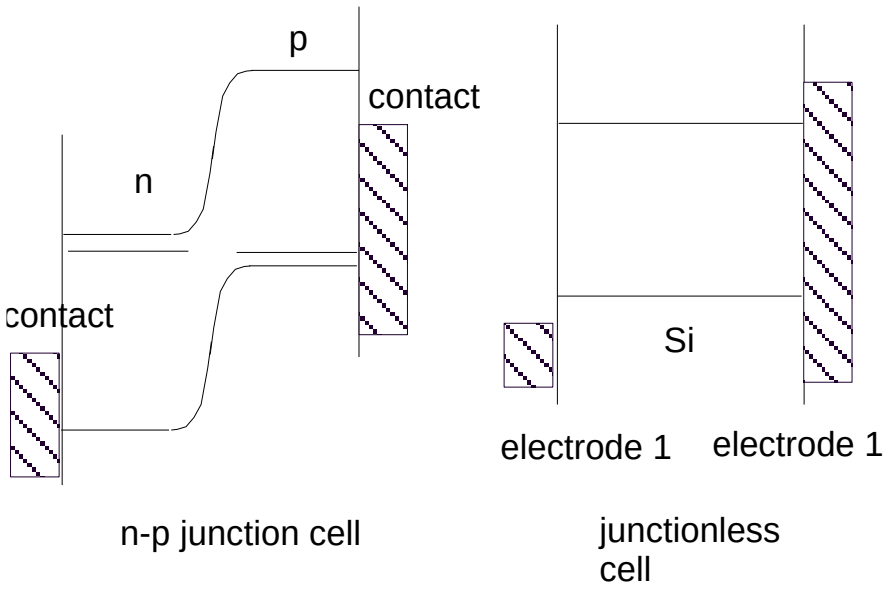


Fig. 4-1 Two designs concepts for Si solar cells.

After some years, many of the problems of OLED electrode materials have now been overcome [8, 9], and there has been much exchange of ideas between the various branches of semiconductor technologies. It is therefore interesting to study some of the material design concepts involved in dopant-free photovoltaics and carrier-selective electrodes.

Another aspect of photovoltaic design has been the effort to maximize cell efficiency by minimizing the surface recombination velocity of carriers. This has motivated an efficient

passivation of surfaces, to remove surface defects [4]. But, generally, it will involve directing the photo-generated electrons and holes to different electrodes so that they avoid each other [10].

For selective electrodes, the idea is to use two electrodes of sufficiently different work functions to give the desired open circuit voltage [4]. In this choice, many workers consider the work functions of the free metals, and then use the electron affinity rule to work out the cell's band diagram [4], Fig 4-2(left), and thus maximum open circuit voltage. However, when two metals are in contact with Si, they are subject to the effects of Fermi level pinning (FLP). This greatly reduces the difference of their *effective* work functions. This pins the Fermi energy of each contact to the charge neutrality level (CNL) of Si [11-14]. This greatly reduces the open circuit output voltage of the cell to near zero. The degree of FLP is given by the pinning factor, $S = d\phi_n/d\Phi$, where ϕ_n is the electron Schottky barrier height (SBH) and Φ_M is the metal work function. S = 1 in the unpinned or Schottky limit, and S = 0 in the pinned or Bardeen limit. For Si, experimentally S ~0.05, so assuming a Φ difference of say 1.4 eV (Pt to Al) [15], that the effective work function difference is now SΦ = 0.07 eV, and the maximum open circuit voltage is only a very small fraction of the Si band gap. This is an extremely large reduction, which is shown in Fig 4-2(right). Thus carrier selective electrodes will give very inefficient cells if used directly on Si.

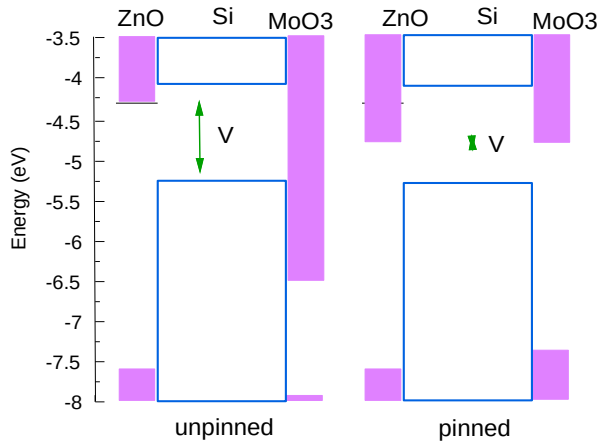


Fig 4-2. Schematic band alignments for (left) idealized case of unpinned electrodes (S= 1), and (right) fully pinned electrodes (S=0), for ZnO and MoO_3 electrodes on Si. Energies referred to vacuum level, or Si band edges.

The only way is to remove the Fermi level pinning effect at the contacts. This can be carried out by inserting a very thin layer of dielectric, which increases S to 0.5 or higher. This method has been shown to be effective in microelectronics [16-23], and will be shown to be effective also in a PV context. This means that the key point about inserting a dielectric layer on the Si is to unpin the Fermi level; it is not to passivate its interface gap states, which is a separate problem. It should also be noted that the purpose of the dielectric layer is more to reduce the effect of the *intrinsic* metal induced gap states (MIGS), rather than passivate the *extrinsic* surface defect states because the MIGS are much more effective than defects at pinning E_F in a narrow gap semiconductor like Si [13]. Here, we carry out density functional calculations on supercell models of the contact interfaces to illustrate what is happening. Finally we combine our results with a numerical modeling result of Agrawal et al [23] to provide a common approach to the problem.

4.2 Methods

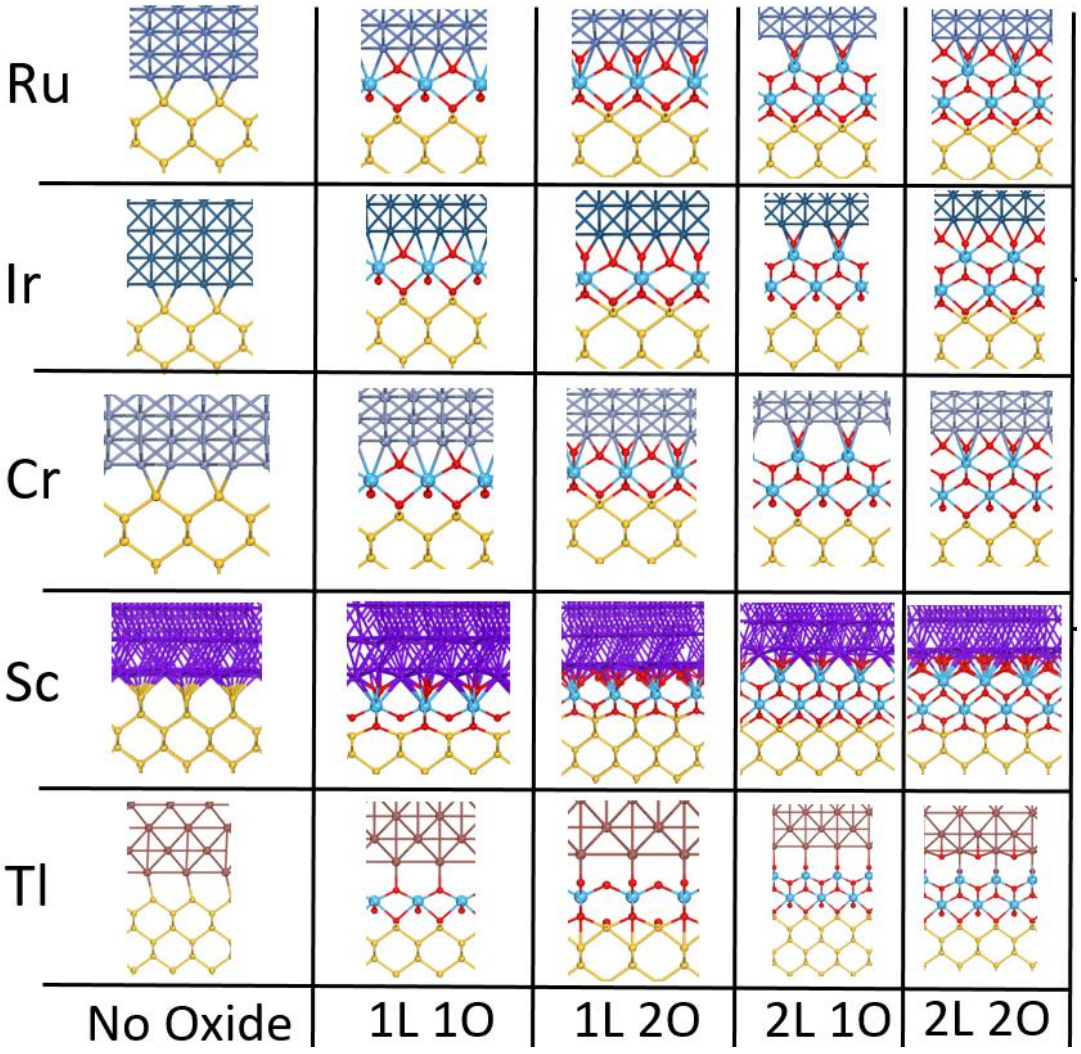
We first discuss the calculation method. We calculate the SBH of various metals on the Si (100) surface using supercells. The plane-wave density functional code CASTEP is used [24], with ultra-soft pseudopotentials, and a plane wave cut off energy of 340 eV. The supercell includes 5-6 layers of metal, 4 unit cells of Si vertically to ensure convergence of the MIGS within the Si layer, and 0-2 unit cell layers of cubic HfO_2. The bottom layer of the Si is terminated by hydrogens, and finally there is 15Å of vacuum gap separating the slabs from their periodic repeats. To keep the total number of atoms not too large, and noting the number of layers in each slab to allow for the convergence of the MIGS, it preferable to have lateral lattice-matching between the various layers where possible, so as not to have an excessive lateral supercell size.

Some metals such as Ru have lattice constants very close to Si, see table 4-1. Other metals such as Ni and Co are reasonably lattice-matched to Si, so that a 1x1 lateral supercell is possible, Fig 4-3. In these cases, the cell size is fixed to that of Si, and the metal lattice is allowed to relax normal to the surface. This is because a metal's work function is primarily a function of its atomic volume and not on individual lattice constant values [25, 26]. For other

metals with a greater lattice mismatch like Ti, the same procedure is used. A larger supercell of 2x2 can be used. However, these metals often show signs of reacting with Si, and so some freezing of atomic coordinates may be necessary. This is particularly the case where oxide insertions are used. Finally, there are cases like Sc where the metal lattice must be rotated with respect to the Si lattice to provide a suitable lattice-matching. Sc is an electropositive metal, and the effects of reactions are reduced by freezing atomic coordinates where necessary.

Table 4-1. In-plane lattice-matching of Si and metals in each supercell.

Transition metal	*Si(100)*	*metals*	*Lattice mismatch (%)*
Sc(0001)	$\sqrt{5} \times \sqrt{5}$	$2\sqrt{2} \times 2\sqrt{2}$	1.89
Tl(bcc)	1×1	1×1	3.03
Ag(fcc)	1×1	1×1	7.69
Ti(fcc)	1×1	1×1	6.55
Cr(fcc)	1×1	1×1	-5.63
Mo(fcc)	1×1	1×1	4.29
Ru(fcc)	1×1	1×1	-0.39
Co(fcc)	1×1	1×1	-8.31
Ni(fcc)	1×1	1×1	-8.65
Ir(fcc)	1×1	1×1	0.93
Pt(fcc)	1×1	1×1	3.44
MoO_3	1×2	1×2	$-2.21, 2.07$



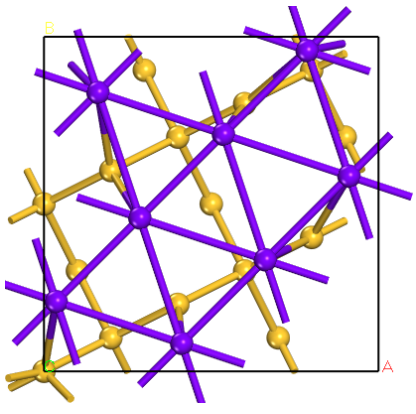


Fig 4-3. Relaxed atomic geometries for five representative metals on Si, and on Si with one or 2 layers of inserted HfO_2, Hf: light blue, O:red, Si:yellow. Bottom shows a top view of the Sc(0001) layer on the top two layer of Si(100).

On the other hand, the high work function oxide MoO_3, a layered material, provides almost van der Waals bonding between itself and the Si surface. Given this, the Si(100) surface is given a 2x1 reconstruction to minimize the number of Si dangling bonds on its surface. In this case, there is a good lattice match and the resulting structure has no direct covalent bonds between the Si and the oxide, as in Fig 4-4. We need to take van der Waals force into account. Therefore, we add the correction term using the Tkatchenko-Scheffler (TS) scheme to obtain correct interface distance and energy.

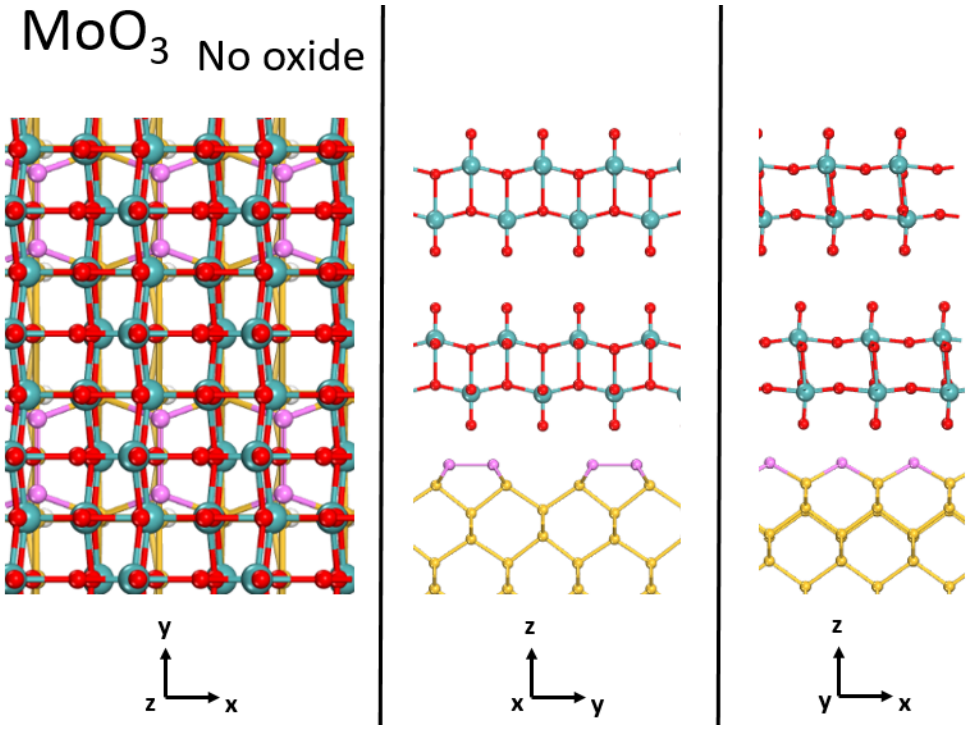


Fig 4-4. Top and side views of MoO_3 on the 2x1 Si(100) surface. Mo=blue/grey, Si = orange, top Si layer = purple. O = red.

We then calculate the atomic structure and the p-type SBH, which is defined as the energy from the Fermi level to the valence band maximum (VBM) of Si. The VBM energy is found from the partial density of states (PDOS) of a Si layer well away from the interface. The band edge is not always clear because the interaction of metals or the oxide states with valence

states of Si. In other cases, we have used a shallow core state such as the Ge 3d level as a reference level, but this is not possible for Si. The VB minimum at -12.5 eV is the closest to a reference state in the Si system.

The plane wave energy cut-off is chosen to be 340eV. A 4x4x1 Monkhorst-Pack k point mesh is used for the 1x1 cell from Tl to Pt. For Sc, a 2x2x1 k point mesh is used. For MoO_3, a 4x2x1 k point mesh is used. All structures are relaxed to a residual force of less than 0.03 eV/Å.

We also study the cases of 1 or 2 monolayers of HfO_2 inserted between the Si and the metal. HfO_2 is chosen as the oxide because it has a cubic phase, and which is reasonably well lattice-matched to Si, when its lattice is rotated by 45 degrees. At each interface, one can retain either one or two oxygens per Hf atom, as shown in Fig 4-3. The Si:HfO_2 interface is always taken to contain two oxygens per Hf. One oxygen gives a metallic interface, while the polar interface [27, 28] with two oxygens per Hf gives an insulating interface because it is a closed shell [29, 30]. This interface reconstructs to give a Si-O-Si / Hf-O-Hf termination, Fig 4-3. On the other hand, the metal / HfO_2 interface is calculated for both the one and two oxygen cases. The bonding at this interface was studied earlier [31, 32]. The interfacial oxygen is most stable in a 4-fold bonded configuration. This site is not tetrahedral, but has two of the metal-O bonds rotated [32], Fig 4-3. Here, the one O case is the non-polar, and the two O case is polar [31]. There is a sizeable interface dipole of ~0.8 eV between these two configurations, for the case of metals on bulk HfO_2 as seen earlier [31].

The size of this dipole for the different interfaces makes the results difficult to analyze. For the HfO_2/metal interfaces, a dipole of 0.8 eV is less than the GGA band gap of HfO_2, ~3.5 eV, so it is possible. But for the Si/HfO_2/metal system, the dipole is larger than the Si GGA band gap of ~0.45 eV. In this case, the SBH is easily affected by the band edge energies. One way to solve this problem is to use hybrid functionals to widen the Si band gap. However, these are relatively slowly converging when the system involves also metals. Instead, we create a modified Si pseudopotential, with a larger GGA band gap of 1.36 eV and higher sigma band of s orbitals of 5.5eV shown in Fig 4-5, to plot the various barrier heights. This is

similar to a modified potential once used for Ge [33]. Note that all structural relaxations use the unmodified Si pseudo-potential.

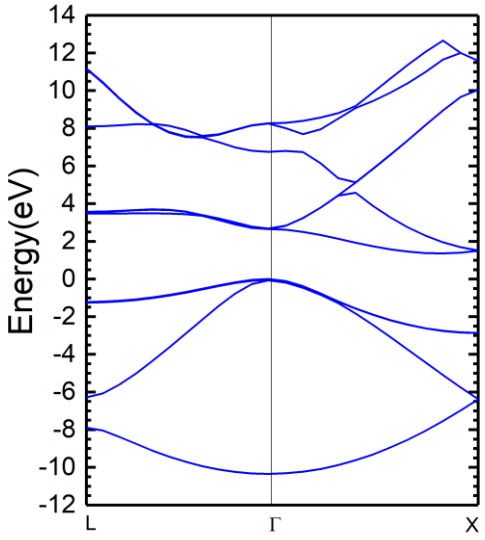


Fig. 4-5 Band Structure of the Si using modified pseudopotential.

4.3 Results

We use 12 metals or conductive metal oxides to cover a wide range of work functions from Sc to MoO_3. The work functions we use are from Michaelson [15]. As chapter 1 illustrated, the n-SBH follows a linear dependence to metal work functions.

$$\phi_n = S(\Phi_M - \mathrm{E}_{cnl}) + E_{cnl} - \chi_s \tag{4.1}$$

Here, Φ_M is the metal work function, E_{cnl} is the charge neutrality level (CNL) of Si and χ_s is the electron affinity of Si. S is the pinning factor.

Fig 4-6 shows the calculated n-type SBH for the metals of Si. They cover a wide range of metal work functions, from Sc to MoO_3. We find a pinning factor of S = 0.03. This value is a little lower than the experimental value. Fig. 4-6 also shows the calculated p-type SBH for the oxide inserted interfaces. There are various points to notice. The data is grouped into

those for one layer of HfO_2 (with one or two O's) and those for 2 layers of HfO_2, with 1 or 2 oxygens next to the metal. We see that the slope S is the same for both O cases for one layer, and both cases for two layers, while there is a substantial upward shift depending on the number of oxygens in the interface. Thus the slope depends on the number of HfO_2 layers. This is correct, it depends on the degree of decay of the MIGS across the oxide layer, while the vertical shift depends on the oxygen-derived dipole at the metal-oxide interface. It is also seen that when S increases beyond 0.5, then the SBHs rapidly hit the band edges of the Si because of the large work function range used in our modeling.

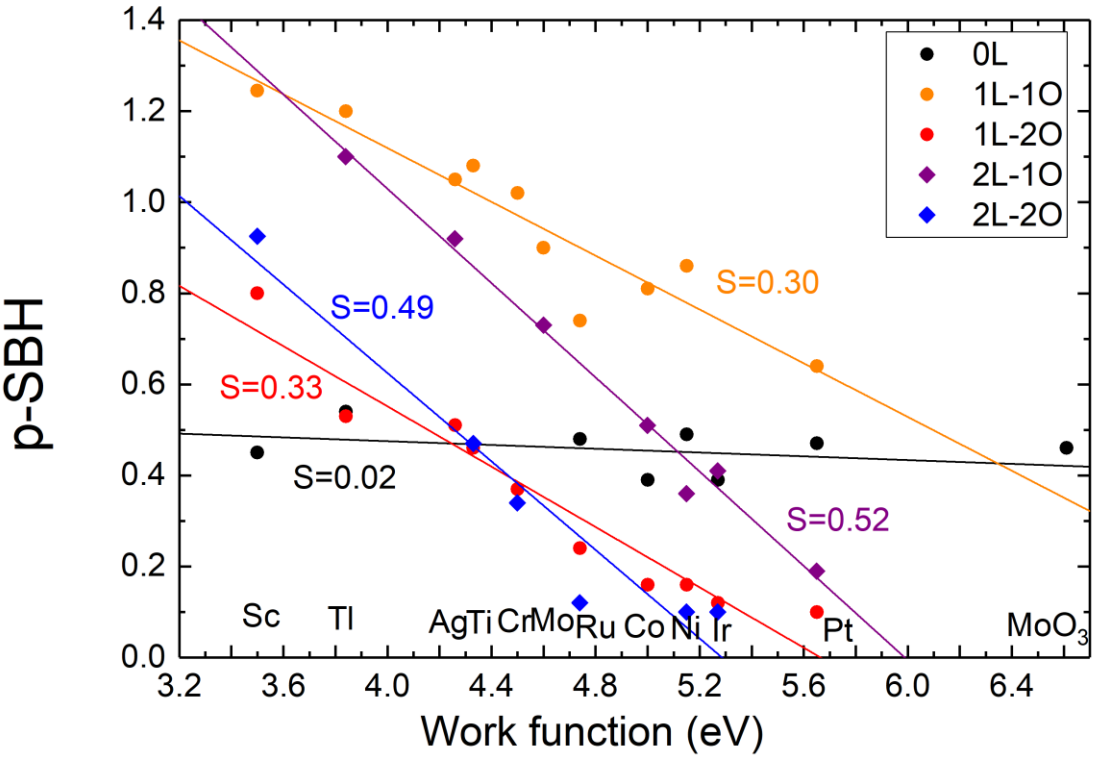


Fig 4-6. Calculated p-type Schottky barrier height (from VBM) vs. experimental work function of metals, after Michaelson [15]). Note Si itself has a very small value of S, and that S value then varies mostly with the number of HfO_2 layers, one or two, irrespective of the oxide termination. The termination creates a dipole (vertical shift in SBH values) between cases of 1 or 2 oxygens. This shift does not affect the decay rate of the MIGS caused by inserting the oxide layer.

We see that the pinning factor S = 0.29 for a HfO_2 monolayer is sufficient to unpin E_F that electrodes of MoO_3 (WF = 6.6 eV) and nano-dot ZnO (WF = 3.5 eV) will give a voltage difference of SΦ = 0.78 V, whereas two layers of HfO_2 will give SΦ = 0.5x2.4 = 1.2 V.

We can now combine our results with those of Agrawal [16]. Table 4-2 gives the band gap, charge neutrality level (CNL), electron affinity (EA), optical dielectric constant, and pinning factor S [11, 12]. S is calculated here using the empirical relationship [13],

$$S = \frac{1}{1 + 0.1(\varepsilon_\infty - 1)^2} \tag{4.2}$$

Generally, Agrawal [23] mostly used values calculated previously [11] with this equation. However, for a few such as ZnS, ZnSe, and GeO_2 they used the method of band structures of imaginary k vector, which gave incompatible values. The CNL is the energy at which the MIGS are filled up to on a neutral surface. We included CNL values calculated elsewhere by us in table 4-2.

Table 4-2. Band gap, charge neutrality level (CNL) from valence band edge, electron affinity (EA, all in eV, optical dielectric constant (given as square of refractive index) and pinning factor S (dimensionless).

	E_{gap}	CNL	EA	ε_∞ (n^2)	S
Si	1.1	0.2	4.05	12	0.05
TiO_2	3.05	2.0	4.0	6.81	0.23
$SrTiO_3$	3.3	2.2	4.1	5.81	0.3
$a\text{-}Al_2O_3$	6.2	2.8	2.5	2.76	0.76
HfO_2	5.8	4.0	2.0	4	0.53
La_2O_3	5.9	2.5	2	4	0.53
Ta_2O_5	4.3	3.3	3.96	4.54	0.44
GeO_2	6.2	3.1	3.2	2.56	0.80
MgO	7.8	3.9	0.8	3.03	0.71
SnO_2	3.6	4.1	4.5	4	0.53
In_2O_3	2.8	3.3	4.5	5.5	0.33
Ga_2O_3	5.0	4.5	3.5	3.7	0.58

ZnO	3.4	3.27	4.35	3.6	0.68
ZnS	3.72	2.1	3.82	5.57	0.32
ZnSe	2.68	1.6	4.09	6.81	0.23
NiO	3.8	1.8	3.0	5.7	0.31
MoO_3	3.0	2.6	6.6	4.4	0.44

We see that the S value calculated here for two layers of inserted HfO_2 is similar to that given by the empirical formula. This is the key result, as this depends on the decay rate of the MIGS across the inserted oxide layers, and this pinning factor calculated by first principles is the same as that found by the empirical relationship.

4.4 Conclusion and discussion

The unpinning effect of the inserted oxide layers is balanced against the higher resistance due to lower tunnelling probability for thicker layers. This is a critical factor in microelectronics where current densities are high. However, in photovoltaics, current densities are much lower, and this is less of a constraint. Interestingly, our results find that two layers of HfO_2 or 1.0 nm is sufficient to unpin E_F. It was found that an optimum thickness of MgO was ~1.0 nm [5]. Table 4-2 shows that MgO has a slightly higher pinning factor of 0.71 than HfO_2, but in a similar range.

The electrodes for the selective contacts are as follows. A suitable anode contact is MoO_3, with slight O deficiency to be n-type [34-36]. There are various possibilities for the cathode. Originally Al/LiF was used in OLEDs. However, this is disfavoured because the Al and LiF must be deposited in one particular order [9] which is inconvenient. It was recently found by Wan et al [5] that Mg-rich MgO layers are a suitable cathode. This is likely to be internally nanostructured, to create the conductive but low work function property. Using this electrode, Wan et al [5] found a minimum of the resistance at 1.0 nm oxide thickness using Al_2O_3. Another candidate is nanostructured Zn silicate, which recently developed and has a work function of ~3.5 eV, much lower than that of uniform ZnO, ~4.2 eV [9, 37].

Reference

1. M A Green, Solar Energy Mats. Solar Cells **143** 190 (2015) PERC
2. J H Yang, S J Kang, Y Hong, K S Lim, Elec. Device Lett. **35** 96 (2013)
3. J Bullock, M Hettick, J Geissbuhler, A J Ong, T Allen, C M Sutter-Fella, T Chen, H Ota, E W Schalter, S DeWolf, C Ballif, A Cuevas, A Javey, Nature Energy **1** 15031 (2016).
4. J Melskens, B W H van de Loo, B Macco, L E Black, S Smit, W M M Kessels, IEEE J Photovoltaics **8** 373 (2018)
5. Y Wan, C Samundsett, J Bullock, M Hettick, T Allen, D Yan, J Peng, Y Wu, J Cui, A Javey A Cuevas, Adv. Energy Mats. **7** 1601863 (2017)
6. J Bullock, Y Wan, Z Xu, S Essig, M Hettick, H Wang, W Ji, M Boccard, A Cuevas, C Ballif, A Javey, ACS Energy Letts **3** 508 (2018)
7. J Cui, Y Wan, Y Cui, Y Chen, P Verlinden, A Cuevas, Appl. Phys. Lett. **110** 021602 (2017)
8. J Huang, A Wan, A Kahn, Mater. Sci. Eng. **64** 1 (2009)
9. H Hosono, J Kim, Y Toda,T Kamiya,S Watanabe, Proc. Nat. Acad. Sci. **114** 233 (2017)
10. U Wurfel, A Cuevas, P Wurfel, IEEE J Photovoltaics **5** 461 (2015)
11. J Robertson, J Vac Sci Technol B **18** 1785 (2000)
12. J Robertson, J Vac Sci Technol A **31** 050821 (2013), J Robertson, S J Clark, Phys Rev B **83** 075205 (2011)
13. W Mönch, Phys. Rev. Lett. **58** 1260 (1987)
14. J Tersoff, Phys. Rev. Lett. **52** 465 (1984)
15. H B Michaelson, J. Appl. Phys. **48** 4729 (1977)
16. T Nishimura, K Kita, A Toriumi, Appl. Phys. Exp. **1** 051406 (2008)
17. M Kobayashi, A Kinoshita, K Saraswat, H S P Wong, Y Nishi, Tech Digest VLSI (2008) p54;
18. M Kobayashi, A Kinoshita, K Saraswat, H S P Wong, Y Nishi, J. Appl. Phys. **105** 0236702 (2009)
19. R R Lieten, S Degroot, M Kuijk, G Borghs, Appl. Phys. Lett. **92** 022106 (2008); GeN
20. R R Lieten, V V Afanasev, N H Thomas, S DeGroote. W Walukiewicz, G Borghs, J Electrochem Soc **158** H358 (2011) GeN
21. L Lin, H Li, J Robertson, Appl. Phys. Lett. **101** 172907 (2012)
22. A M Roy, J Lin, K C Saraswat, IEEE ED Letts. **31** 1077 (2010) anal
23. A Agrawal, N Shukla, K Ahmed, S Datta, Appl. Phys. Lett. **101** 042108 (2012) anal
24. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, Zeitschrift Für Krist. **220**, 567 (2005)
25. N Lang and W Kohn, Phys. Rev. B **3** 1215 (1971).
26. A Kiejna and K F Wojciechowski, Metal Surface Electron Physics. Elsevier (1996).
27. P W Tasker, J. Phys. C **12** 4977 (1979)
28. J Goniakoweski, C Noguera, Phys. Rev. B **66** 085417 (2002)
29. P W Peacock, J Robertson, Phys. Rev. Lett. **92** 057601 (2004)
30. P W Peacock, K Xiong, K Tse, J Robertson, Phys. Rev. B **73** 075328 (2006)
31. K Y Tse, D Liu, J Robertson, Phys. Rev. B **81** 035325 (2010)
32. K Y Tse, J Robertson, Phys. Rev. Lett. **99** 086805 (2007)

33. M Houssa, V Afanasev, A Stesmans, G Pourtois, M Meuris, M Heyns, Appl. Phys. Lett. **95** 162109 (2009)
34. M Kroger, S Hamwi, J Meyer, T Riedl, W Kowalsky, A Kahn, Appl. Phys. Lett. **95** 123301 (2009)
35. Y Guo, J Robertson, Appl. Phys. Lett. **105** 222110 (2014)
36. C Battaglia, S M deNicolas, S DeWolf, X Yin, M Zheng, C Baliff, A Javey, Appl. Phys. Lett. **104** 113902 (2014)
37. N Nakamura, J Kim, H Hosono, Adv. Electron Mater. **4** 1700352 (2018)

Chapter 5 Charge transfer doping of Graphene without degrading carrier mobility

Density functional calculations are used to analyse the charge transfer doping mechanism by molecules absorbed onto graphene. Typical dopants studied are $FeCl_3$, $AuCl_3$, SbF_5, HNO_3, MoO_3, Cs_2O, O_2 and OH. The Fermi level shifts are correlated with the electron affinity or ionization potential of the dopants. We pay particular attention to whether the dopants form direct chemisorptive bonds which cause the underlying carbon atoms to pucker to form sp^3 sites as these interrupt the π bonding of the basal plane, and cause carrier scattering and thus degrade the carrier mobility. Most species even those with high or low electronegativity do not cause puckering. In contrast, reactive radicals like –OH cause puckering of the basal plane, creating sp^3 sites which degrade mobility.

5.1 Background

Graphene is a two-dimensional material with a unique band structure with bands crossing at the Dirac point [1]. This gives graphene a very high carrier mobility, but the carrier concentration is small so that its overall electrical conductivity is rather low [2]. Thus, graphene must be doped to increase carrier concentration in order to realize some of its main applications such as a transparent electrode in displays or photovoltaic devices [2-6] or as a sensor [7-9]. However, the doping should not degrade its mobility by, for example, introducing Coulombic scattering centres. These would reduce the mobility μ according to μ=a/N dependence [10], where N is the density of centres. This could lead to no net increase in conductivity in an extreme case. Nor should doping interfere with uniform π bonding of the graphene sheet by converting sp^2 to sp^3.

The conventional way to dope a 3-dimensionally bonded semiconductor would be by substitutional doping. This has indeed been carried out for graphene using nitrogen or boron doping [11-13]. Substitutional sites have the advantage in being fully bonded into the lattice

and are thus stable. However, nitrogen can enter the graphene lattice in various configurations, only one of which is the actual doping configuration [13, 14]. The other configurations not only do not dope, they also introduce defects [15, 16] which cause carrier scattering. This 'functionalization' is useful in other contexts such as creating catalytic sites on carbon nanotubes [15]. On the other hand, for graphene as an electrode, it is useful to consider interstitial or transfer doping by physisorbed species [17-24]. These can dope the graphene n-type or p-type, without necessarily creating defects. Transfer doping is also useful to increase the conductivity of contacts, as the high resistance of contacts to graphene in devices can limit the device performance. The transfer doping method is also relevant to doping of other 2D systems like MoS_2 [25] and is frequently used in organic electronics [26].

However, a critical factor not previously studied is whether the dopant forms a weak physisorptive bond or strong chemisorptive bond to the graphene. For the first case, this will allow charge transfer (Fig. 5-1), without modifying the π bonding of the graphene layer and so it should maintain the mobility of graphene. On the other hand, if a short chemisorptive bond is formed, this will convert the underlying C sp^2 to sp^3, so removing the π orbital of that site and degrading the graphene mobility.

Here, we study the charge transfer doping caused by a range of dopants. Some of these were previously used in the intercalation of graphite [27, 28], or the charge transfer doping of organic molecules such as in organic light emitting diodes [29, 30]. It turns out that some of the dopants have very large electronegativities compared to elemental metals, or are strongly electropositive. Interestingly, we find that even strongly electronegative or electropositive species need not form chemisorptive bonds and so are good transfer dopants.

5.2 Methods

The calculations are carried out using periodic supercell models of the graphene and the dopant species using the CASTEP plane-wave density functional theory (DFT) code [31], with ultra-soft pseudopotentials and the Perdew-Burke-Ernzerhof (PBE) form of the generalized gradient approximation (GGA) for the electronic exchange-correlation functional. For an open shell magnetic system $FeCl_3$, we use the GGA+U method, with an on-site potential U of 7 eV

applied to the Fe 3d states. The screened exchange hybrid functional [32] is also used to correct GGA band gap error in the Cs_2O system.

The dispersion correction to the GGA treatment of the van der Waals interaction is included using the Tkatchenko-Scheffler (TS) version [33] of the Grimme scheme [34]. To overcome the error induced by periodical mirror charge, self-consistent dipole correction is implemented. The plane wave cut-off energy is 380eV, as the cut-off energy of oxygen.

For the graphene plus dopant system, a layer-by-layer stacked supercell is created in each case, with a close degree of lattice matching between the graphene and the dopant. A 30Å vacuum layer is included where a vacuum layer is needed. The size of the supercell is given in Table 5-1. A dense 9 × 9 k-point mesh is used to calculate the density of states (DOS), due to the small density of states of Graphene close to the Dirac point. The calculated lattice constant of graphene in PBE is 2.47Å, 0.4% less than the experimental value [1]. The physisorptive binding energy, relevant bond lengths, and any puckering of the graphene sites below the dopant site are given in Table 5-2.

Table 5-1. Supercell and lattice match of Graphene and dopant. M in the mismatch column refers to a molecular dopant where there is no mismatch.

	Graphene supercell/ Dopant supercell	**Mismatch rate (%)**
SbF_5	$\sqrt{3} \times \sqrt{3}/1 \times 1$	1.66
$FeCl_3$	$\sqrt{7} \times \sqrt{7}/1 \times 1$	1.42
$AuCl_3$	$4 \times 4/1 \times 1$	M
MoO_3	$3 \times \sqrt{3}/2 \times 1$	0.34, 7.92
Cs_2O	$\sqrt{3} \times \sqrt{3}/1 \times 1$	1.62
Cl_2	5×5	M
O_2	5×5	M
OH	5×5	M
HNO_3	5×5	M

Table 5-2. Atomic distance, bond length and puckering of Graphene.

	Bond type	Bond length (Å)	Surface distance (Å)	Puckering (Å)	Binding energy (eV)
OH	O-H	0.98	–	0.51	-1.64
	C-O	1.51	–		
O_2	O-O(in O_2)	1.24	3.29	0.09	-0.13
HNO_3	O-H(in H_2O)	0.98	3.28	0.06	-0.39
	N-O(in NO_2)	1.23	2.60		
	N-O(in NO_3)	1.27	3.25		

Doping causes a shift in the system's Fermi energy away from the Dirac point of the graphene, as in Fig. 5-1. This shift is compared to the Fermi energy, ionization potential (IP) or electron affinity (EA) of the isolated dopant system. These energies are calculated using a periodic supercell of the dopant species plus vacuum gap. The electrostatic potential is calculated for the dopant system layers, averaged along the layers. The potential in the vacuum gap region gives the vacuum potential. The energy of the valence band maximum is then compared to the vacuum energy to give the ionization potential, and with the band gap, the electron affinity.

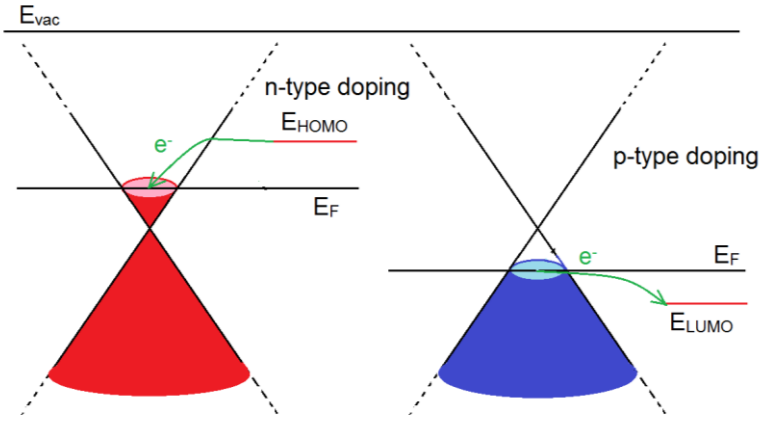


Fig 5-1. Schematic of n-type and p-type doping process in Graphene.

5.3 Results

5.3.1 $AuCl_3$

We first consider Lewis acids such as $AuCl_3$ and $FeCl_3$. $FeCl_3$ has been more heavily studied, but $AuCl_3$ is simpler computationally because it does not involve open shell d electrons. Crystalline $AuCl_3$ consists of stacked layers of Au_2Cl_6 molecular units. The Au_2Cl_6 molecule consists of two planar edge-connected $AuCl_4$ squares. The supercell consists of alternate graphene and $AuCl_3$ layers along the z-axis. Fig. 5-2(a) shows the 4x4 graphene supercell with the planar Au_2Cl_6 units separated from each other in-plane at a similar distance as in pure $AuCl_3$. The position of Au_2Cl_6 units on the graphene is allowed to vary to minimize the total energy.

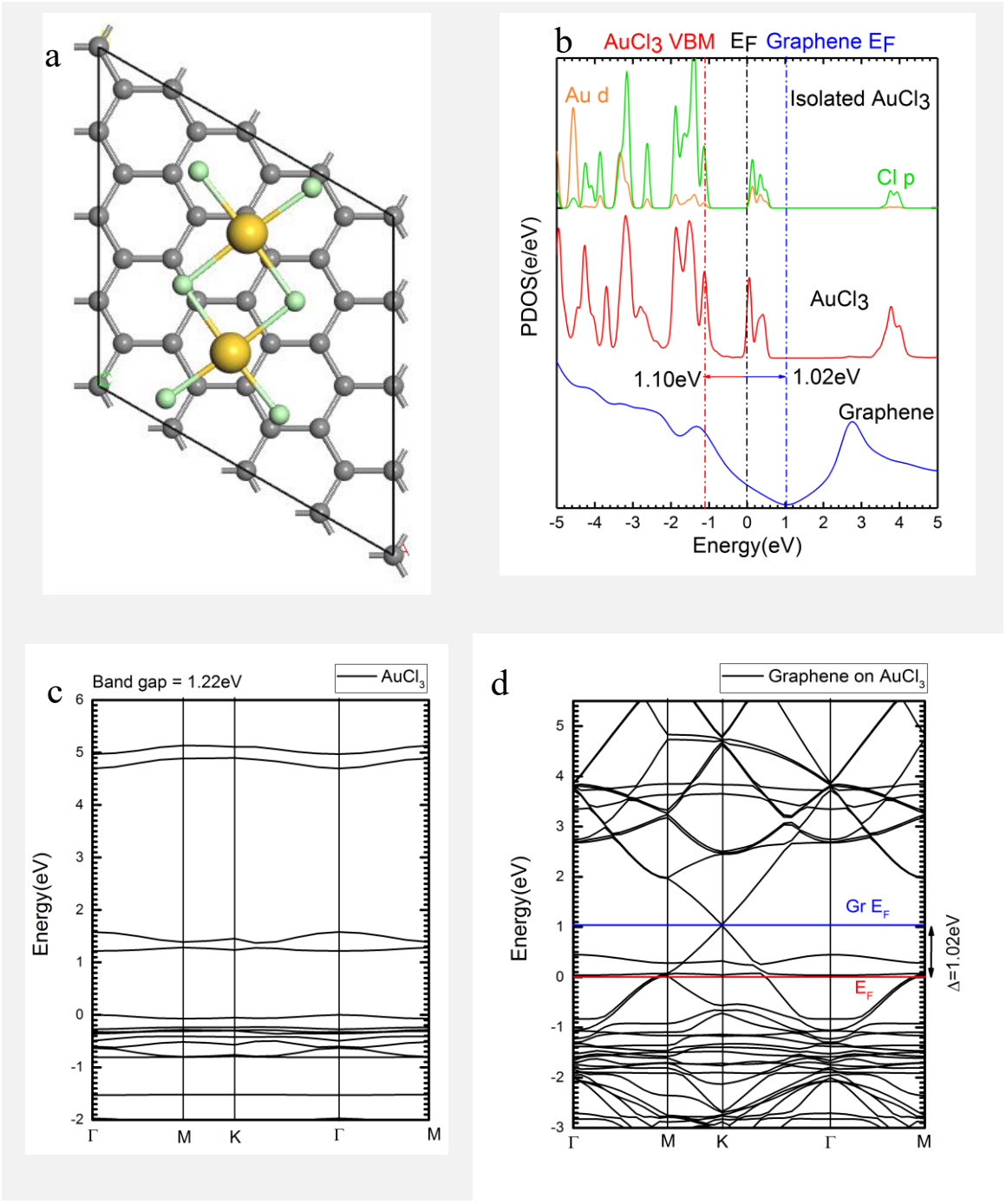


Fig 5-2. (a) Au_2Cl_6 molecular on 4x4 Graphene supercell. (b) PDOS of isolated $AuCl_3$ and $AuCl_3$/Graphene system. (c) Band Structure of isolated pure Au_2Cl_6 in the hexagonal lattice. (d) Band Structure of the combined system.

Fig 5-2(c) shows the band structure of isolated pure Au_2Cl_6 in the hexagonal lattice. The Au_2Cl_6 is a semiconductor with a band gap of 1.22 eV. The Au 5d band is filled to $d^{9.6}$. The conduction band consists of the Au s state and Cl p states. Fig 5-2(d) shows the band structure of the combined system. As a 4 × 4 supercell was used, the graphene Dirac point still lies at K and can be recognized as the crossed bands at 1.02 eV. This shows that the shift of the Fermi energy E_F due to this $AuCl_3$ doping concentration is 1.02 eV.

Fig 5-2(b) shows these results in a density of states (DOS) plot. The doping has occurred by a transfer of electrons from the graphene valence band into the $AuCl_3$ conduction band, filling its conduction bands lying just below 0 eV in the central panel of Fig 5-2(b). (If any Cl vacancies form, they are shallow donors, and these would also become filled by transfer doping.) The carbons of the graphene lattice are found to maintain their planar geometry and do not buckle. The dopant-C separation is 3.35Å (Table 5-3), so the bond is weak and physisorptive, and no puckering of the underlying C site occurs. This will cause no reduction in mobility.

5.3.2 $FeCl_3$

We next consider $FeCl_3$, which is also a Lewis acid like $AuCl_3$. It has been used extensively as an intercalant of graphite [37-43], as discussed by Li and Yue [43]. Solid $FeCl_3$ forms a layered system of Fe_2Cl_6 edge-connected octahedral connected along three directions at 120° to each other. The Cl sites are rotated slightly off the vertical. A hexagonal supercell lattice of graphene and $FeCl_3$ can be made with a 23Å periodicity [37]. On the other hand, we create a smaller, more efficient √7x√7 supercell using a 1x1 periodicity of the $FeCl_3$ sublattice and a √7x√7 periodicity of the graphene, as in Fig 5-3(a). $FeCl_3$ is a magnetic semiconductor with a 0.7 eV band gap. A vertical stacking of one $FeCl_3$ layer and one graphene layer along Oz is ferromagnetic. A stacking of two $FeCl_3$ layers and two graphene layers along Oz, as here, allows the $FeCl_3$ to be anti-ferromagnetically (AF) ordered, which simplifies the band structure plots (the spin-up and spin-down bands are degenerate). Fig 5-3(c) shows the AF bands of isolated $FeCl_3$ calculated for a value of U= 7 eV, with the 0.7 eV band gap. The Fe 3d occupancy is $d^{5.6}$.

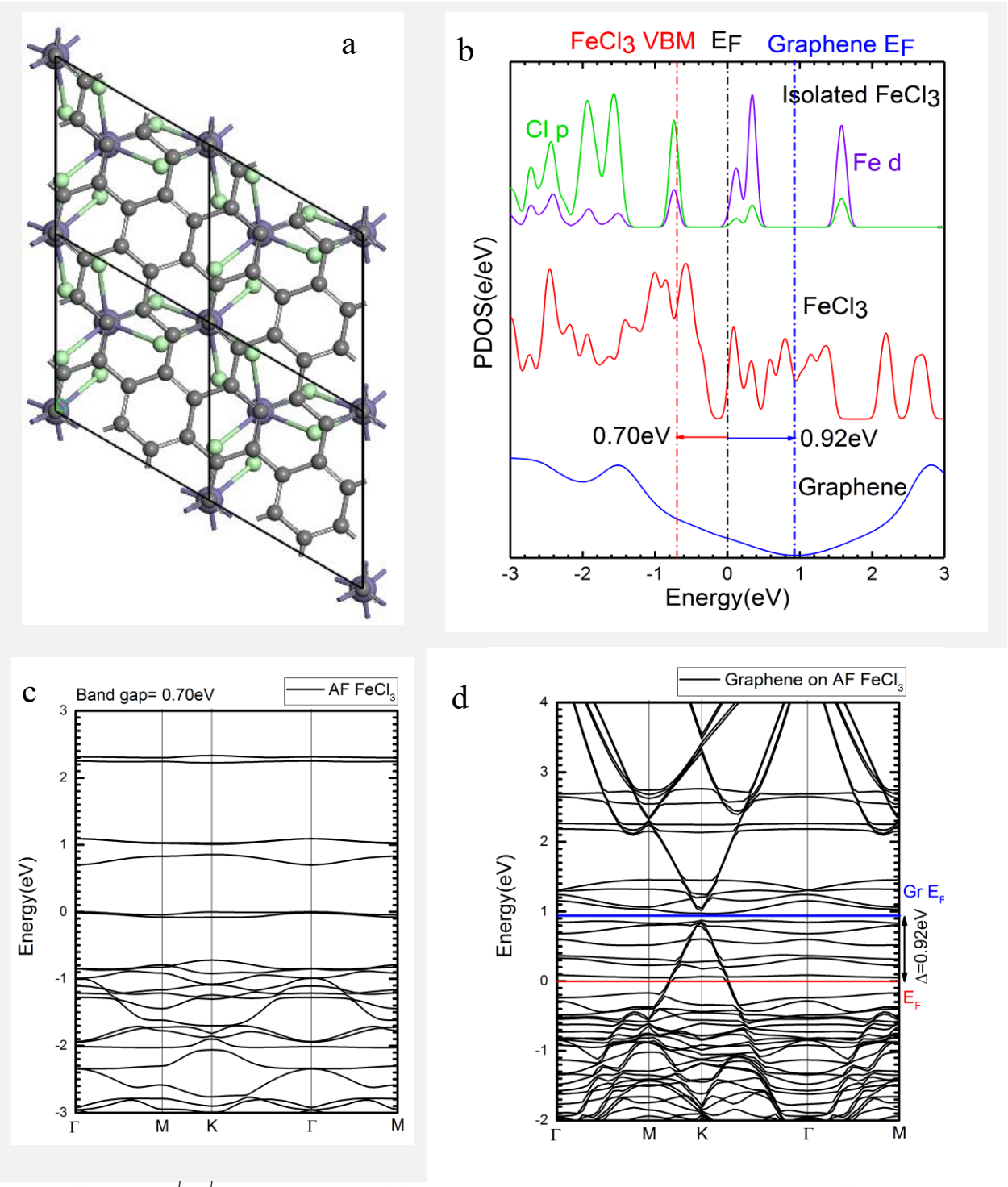


Fig 5-3. (a) $FeCl_3$ on $\sqrt{7}x\sqrt{7}$ Graphene supercell. (b) PDOS of isolated AF $FeCl_3$ and $FeCl_3$/Graphene system. (c) Band Structure of isolated pure AF $FeCl_3$. (d) Band Structure of the combined system.

Fig 5-3(d) shows the band structure of the combined system. The graphene Dirac point can be recognized at the K point 1.0 eV above the Fermi energy. Fig. 5-3(b) shows the density of states for the combined system and for the isolated $FeCl_3$. Doping has occurred by transfer of electrons from the upper graphene valence band to the $FeCl_3$ conduction states at -0.1 eV in Fig 5-3(b).

As for $AuCl_3$, $FeCl_3$ forms a long physisorptive bond of 3.54Å to the graphene. No puckering of underlying carbon occurs, so the transfer doping of graphene by $FeCl_3$ does not degrade its mobility.

5.3.3 SbF_5

We next consider the strongest Lewis acid, SbF_5. Condensed SbF_5 can be considered to form a network of corner-sharing octahedral with the F sites vertically above each other. The SbF_5 units form chains which are conveniently lattice-matched to graphene, when a supercell of 1x1 SbF_5 and $\sqrt{3}$x$\sqrt{3}$ is used, as in Fig 5-4(a).

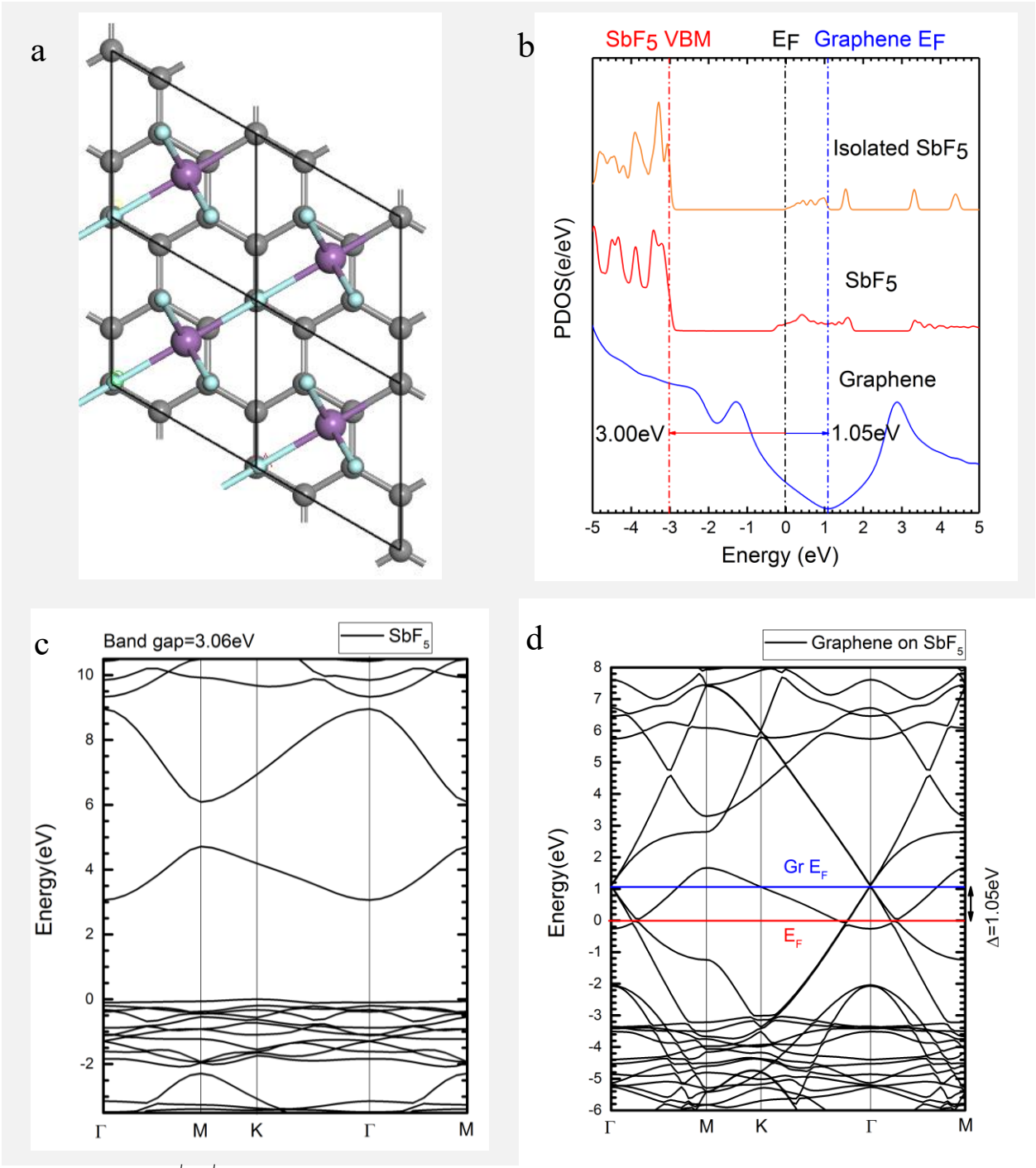


Fig 5-4. (a) SbF_5 on $\sqrt{3}x\sqrt{3}$ Graphene supercell. (b) PDOS of isolated SbF_5 and SbF_5/graphene system. (c) Band Structure of isolated pure SbF_5 single layer in the hexagonal lattice. (d) Band Structure of the combined system.

Fig 5-4(c) shows the band structure of isolated SbF_5 in the unit cell of Fig 5-4(a). It is a semiconductor with a GGA band gap of 3.06 eV and a direct gap at Γ. This system contains only s,p electrons and Sb is in its +5 valence state. The top of the valence band consists of F 2pπ states the conduction band minimum consists of empty Sb 5s states. The high electronegativity of F accounts for the large ionization potential of SbF_5 of 11eV (table 5-3).

Fig 5-4(d) shows the band structure of the combined system. Due to the orientation of the graphene and SbF_5 sublattices, the Dirac point folds over to appear at Γ, at about 1.0 eV above the combined Fermi energy. Fig 5-4(b) shows the density of states of the combined systems and the isolated graphene. Doping has occurred by transfer of electrons from the graphene valence band into the SbF_5 conduction band. This has caused a 3.0 eV shift of the SbF_5 bands, but only a 1.05 eV downward shift of the E_F of graphene.

Table 5-3 gives the calculated electron affinity, band gap and ionization potential of these compounds. As ideal isolated semiconductors, their Fermi energies would appear near midgap. In practice, the anion vacancy is the lowest energy defect in these systems, and this defect is calculated to be shallow. Thus, in practice, their Fermi energy is likely to lie close to their conduction band edges. The large electronegativity of the halogens means that the valence bands of these systems are very deep below the vacuum level. Even with E_F lying at their conduction band edges, their work functions are still very large, much larger than that of the most electropositive metal, Pt.

Table 5-3. Calculated layer distance, work function, ionization potential and Fermi level shift (FLS) from GGA.

	Layer distance (Å)	**Work Function (eV)**	**Ionization potential (eV)**	**FLS (eV)**
SbF_5	3.65	7.04	10.10	-1.05
$FeCl_3$	3.54	6.97	7.12	-0.92
MoO_3	2.95	6.61	8.64	-0.63
$AuCl_3$	3.35	5.94	7.16	-1.02
Cs_2O	3.75	0.90	2.35	0.95

5.3.4 MoO_3

We now pass to the case of MoO_3. This oxide has been widely used as a p-type dopant and electrode material in organic electronics [29, 30], and has recently been used for p-type transfer doping carbon nanotubes, graphene [21, 22] and MoS_2 contacts [44, 45]. MoO_3 has two forms, the molecule Mo_3O_9, and a layered solid form MoO_3. MoO_3 was previously calculated to have a band gap of 3.0 eV and an electron affinity of 6.61 eV [46]. Its oxygen vacancies were calculated to be shallow. The doping of MoS_2 and carbon nanotubes by MoO_3 layers has already been studied theoretically [21, 45].

An orthorhombic supercell of graphene and MoO_3 was constructed as in Fig 5-5(a). The electronic structure of the combined system was calculated. The large work function of MoO_3, 2 eV below that of graphene, means that there is a strong transfer doping. It is found that the Fermi energy of the combined system has shifted downwards in the graphene by 0.63 eV. In this case, doping has occurred by the transfer of electrons from the graphene valence band to the MoO_3 conduction band states. Nevertheless, the bonds between graphene and the outer O layer of MoO_3 are only physisorptive with a bond length of 2.5Å. MoO_3 does not cause any puckering of the graphene sp^2 sites and thus does not affect the π bonding of the graphene layer. Thus, the C atoms do not act as defects under this doping process. There will be no Raman D peak and no carrier scattering. This is consistent with experiment where notably Chen *et al.* [17] find that MoO_3 doped graphene retains the ability to show a quantum Hall effect, indicating a high carrier mobility.

MoO_3 is a very valuable dopant of graphene because it is a stable dopant, it raises the carrier concentration, it does not degrade the carrier mobility by causing defects, it has a wide band gap so that it is also optically transparent, a very useful combination useful for optical devices [18].

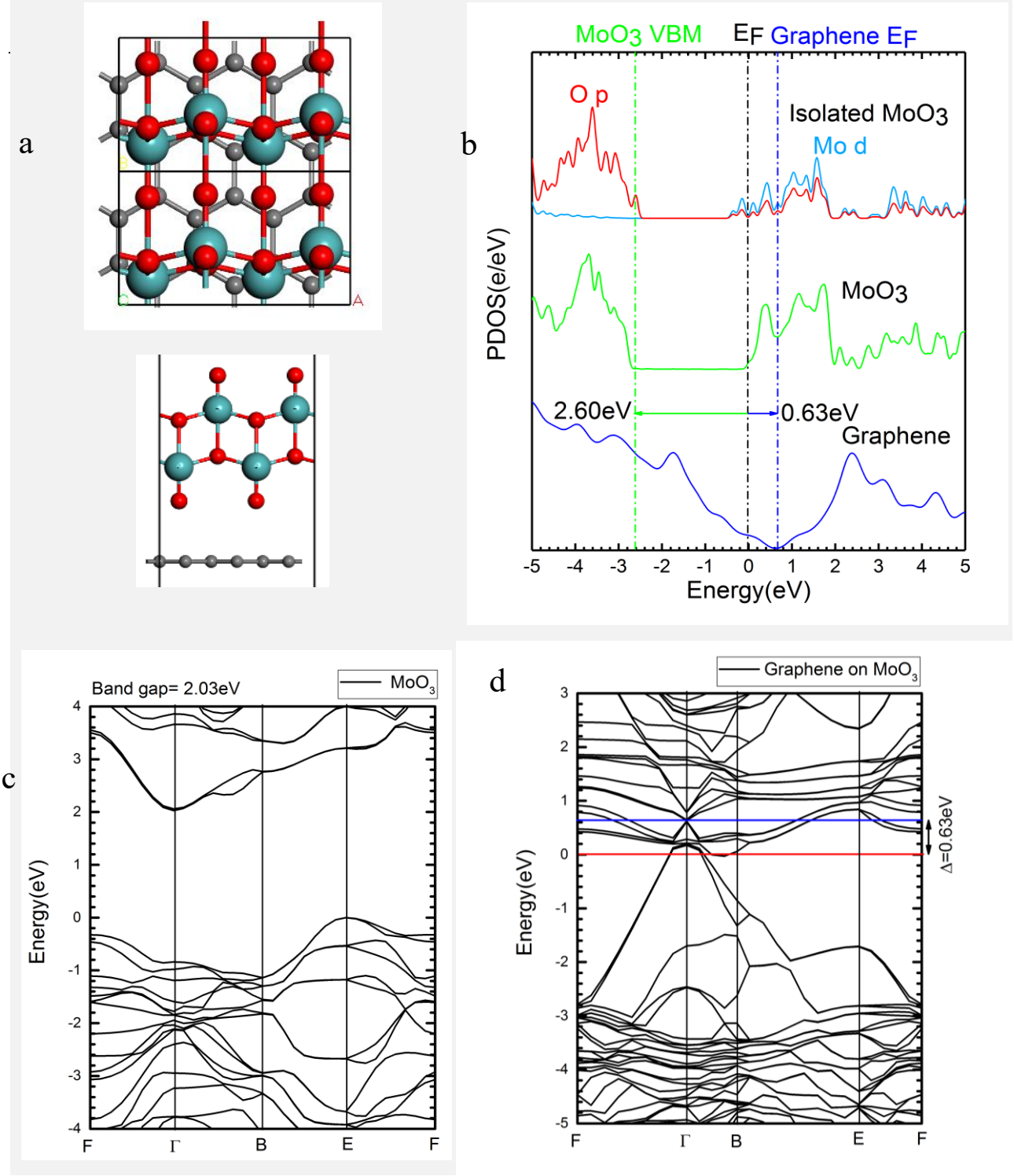


Fig 5-5. (a) MoO_3 on 3x√3 Graphene supercell. (b) PDOS of isolated MoO_3 and MoO_3/Graphene system. (c) Band Structure of isolated pure single layer MoO_3. (d) Band Structure of the combined system.

5.3.5 Cs_2O

We now consider an n-type transfer dopant, CsO_x. Cs carbonate is widely used as an n-type dopant in the organic light emitting diodes, and also can be used to dope graphene [19]. The carbonate precursor dissociates on heating to leave a Cs oxide, which may actually be a sub-oxide. We consider the oxide to be Cs_2O. This has the inverse $CdCl_2$ hexagonal layered structure, with the Cs layers on the outside and O atoms on the inside. Note that whereas the interlayer bonding in $CdCl_2$ is van der Waals, the Cs-Cs bonding in Cs_2O is essentially metallic, not van der Waals. The hexagonal layers are reasonably lattice-matched to those of graphene, with a 1.6% mismatch, as shown in Table 5-1 and Fig 5-6(a). The Cs and O sites lie over the hollow sites of the graphene lattice.

Fig 5-6 (b) shows the band structure of isolated Cs_2O. Cs_2O is a semiconductor with a band gap of 1.44 eV in screened exchange [32] and a very low electron affinity. Its valence band consists of oxygen 2p states. The valence band is very narrow because the O sites are far apart, so the O-O interface controlling the VB width is weak.

Fig 5-6 (c) shows the density of states for the combined system. There is strong n-type doping, with electrons transferred from the Cs_2O valence band into the graphene conduction band. The E_F of graphene is shifted upwards by 0.95 eV by the Cs_2O layer. Nevertheless, the Cs-C bond is long and physisorptive. It is not van der Waals, and no van der Waals correction to GGA is used in this case. The graphene atoms remain unpuckered below the Cs_2O and sp^2 bonding is maintained in the graphene. This behaviour is similar to the behaviour of Cs_2O as an n-type transfer dopant in organic semiconductors [18].

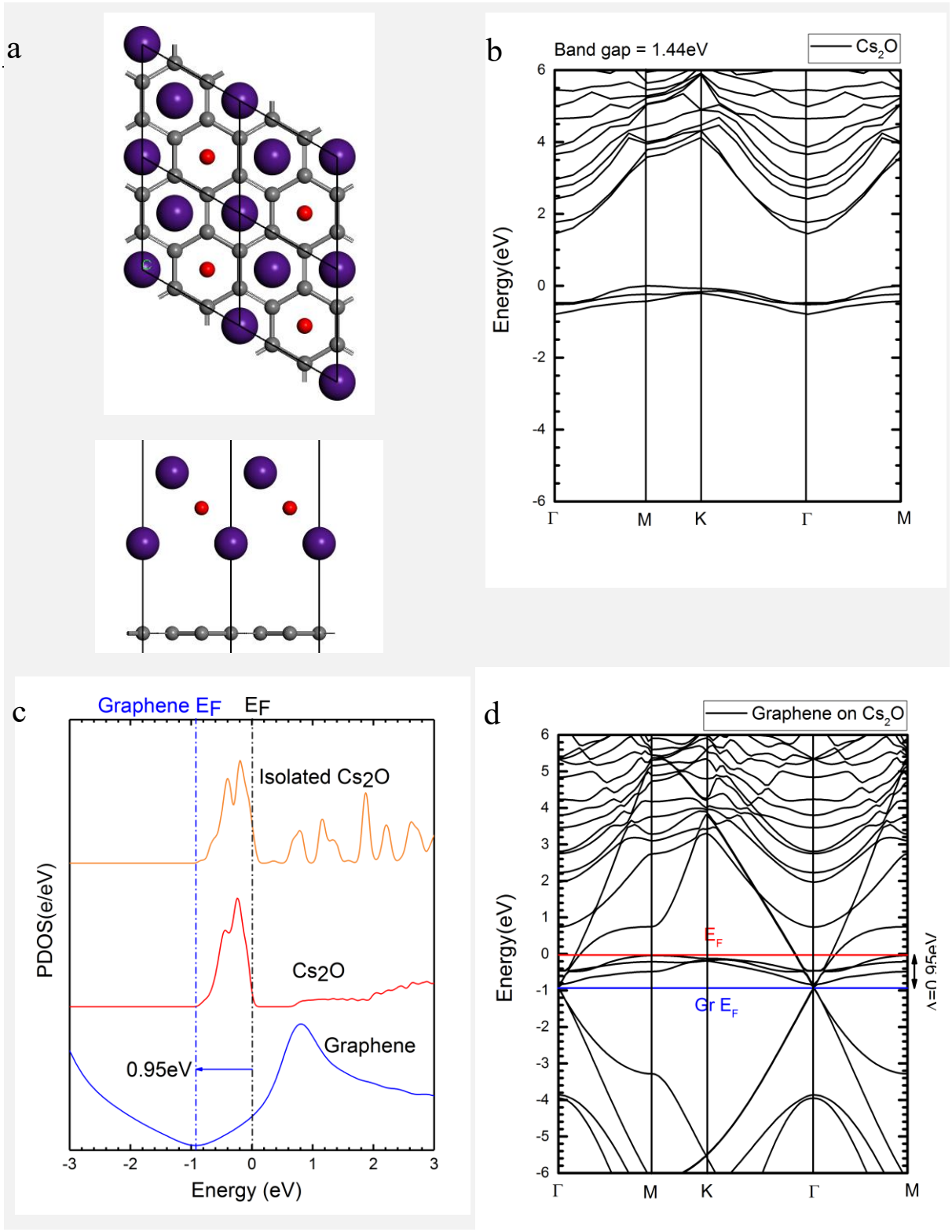


Fig 5-6. (a) Top view and side view of Cs_2O on $\sqrt{3}x\sqrt{3}$ Graphene supercell. (b) Screened exchange band structure of isolated pure single layer Cs_2O. (c) PDOS of isolated Cs_2O and Cs_2O/Graphene system. (d) Band Structure of the combined system.

5.3.6 HNO_3

Nitric acid is another p-type dopant, but it functions differently. Nistor *et al.* [47] studied the absorption of HNO_3 on the graphene surface. They found that HNO_3 dissociates into a NO_3 radical, a NO_2 radical and a water molecule,

$$2HNO_3 \rightarrow H_2O + NO_2 + NO_3$$

HNO_3 is introduced into the 5x5 supercell. Dissociation occurs. These species are allowed to rotate to maximize their stability. The final geometry is shown in Fig 5-7(a). The NO_3 radical lies planar parallel to the graphene plane, with each of its atoms lying on top of a carbon atom. The NO_2 radical and the water molecule lie in a plane normal to the graphene plane, with the central N atom of NO_2 and central O atom of H_2O nose down towards the graphene, as in Fig 5-7(b). These species are physisorbed onto the graphene, and the bond lengths are quite large as expected for physisorption (Table 5-2). The water species causes a very weak buckling of the underlying graphene layer, table 5-2. The binding energy of each species to the graphene is relatively small.

Whereas H_2O is a closed shell system, both NO_3 and NO_2 are radicals each with a half-filled orbital. Critically, the work function of this orbital is greater than that of the graphene, and the state lies deeper below the vacuum level than the Fermi energy E_F of graphene. Thus, they give a single empty state lying below E_F. This leads to an electron transfer from the graphene into the two NO_x species, filling their states, and causing a hole doping of the graphene. As the bond length is long, there is only partial charge transfer. The charge transfer is calculated to be -0.3e on the NO_3 and -0.25e on the NO_2. This lowers E_F of the graphene to -1.10 eV, as shown in Fig 5-7(c). The retention of planar sp^2 bonding in the C sites under the NO_3 and NO_2 physisorbed species means that this does not constitute a defect, there is no Raman D peak and no carrier scattering. This is consistent with experiment. L'Arsie [20] finds no change in the D peak intensity experimentally.

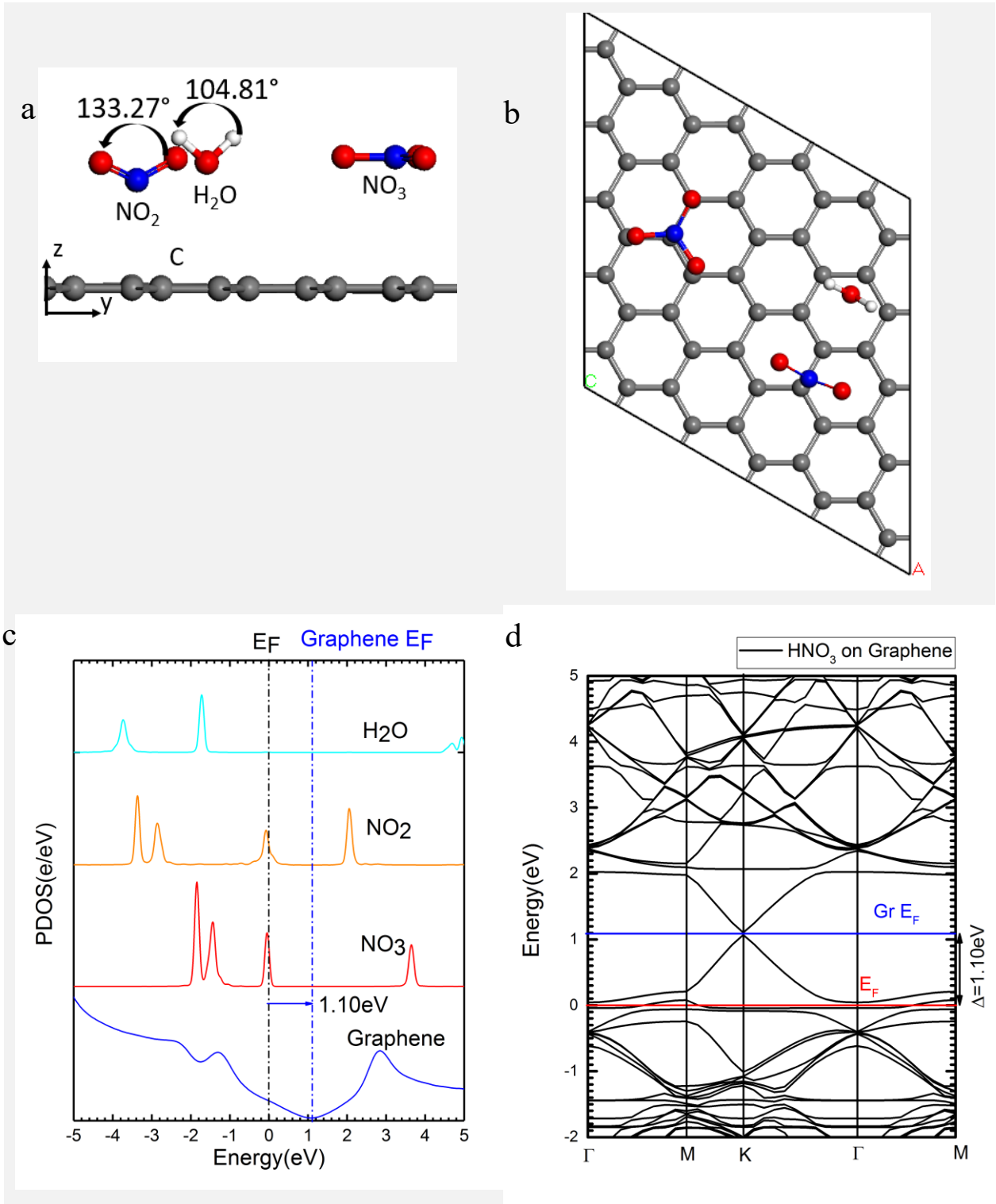


Fig 5-7. (a) Side view and (b) top view of $2HNO_3$ dissolve onto 5x5 Graphene supercell. (c) PDOS of the $2HNO_3$/Graphene system. (d) Band Structure of the combined system.

5.3.7 Cl_2, O_2 and OH radical

We now consider Cl_2. Cl_2 is a closed-shell molecule with a single Cl-Cl bond. It has a filled p_σ state at -12 eV, two filled p_π states and two filled p_π* states, followed by an empty σ* state above its E_F. The Cl_2 molecule is physisorbed onto graphene, but it does not produce doping because it has no empty states below the E_F of graphene [Fig 5-8 (b)]. There is no doping because the empty σ* state is high in energy despite the electronegativity of Cl.

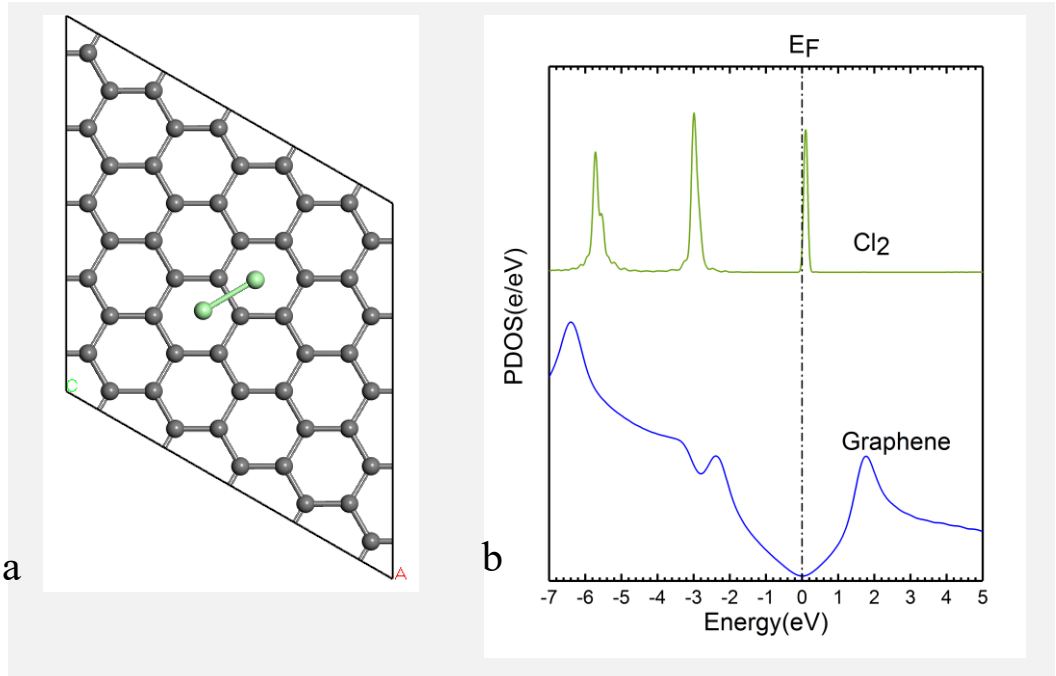


Fig 5-8. (a) Top view of Cl_2 on 5x5 Graphene supercell. (b) PDOS of Cl_2/Graphene system.

Following Cl_2, we consider the O_2 molecule. This molecule is calculated to physisorb in a configuration across a C-C bond, as in Fig 5-9(a). Now, the O_2 molecule is geometrically the same as the Cl_2 molecule, but as its valence is lower, its π* states would be half-filled in the spin unpolarised condition. This configuration is unstable to symmetry breaking to open up a band gap. This occurs by an antiferromagnetic ordering of the σ* spins, with the up-spin states lying below E_F and the down-spins lying above the gap. For the combined O_2 on the graphene system, the gap is small enough that the empty spin-down σ* state lies below E_F of isolated graphene, so there is a sizable charge transfer doping of the graphene by O_2, as shown in Fig 5-9(b). The C-O in this case is long (3.29Å) and physisorptive.

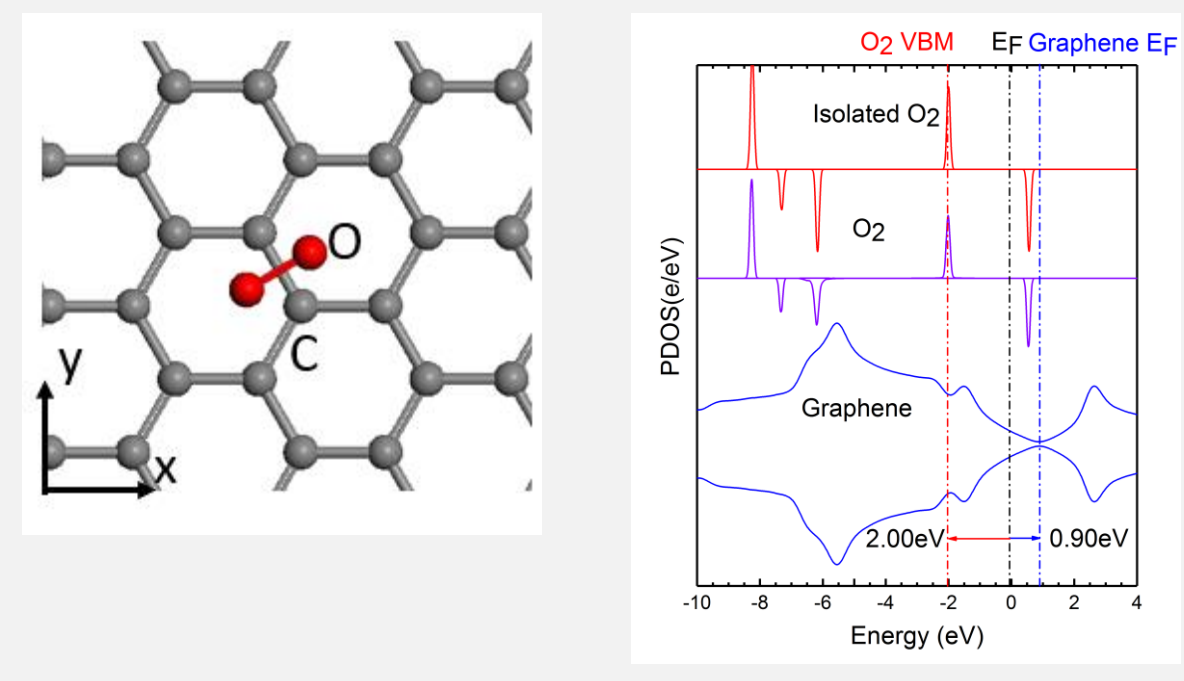


Fig 5-9. (a) Top view of triplet O_2 on 5x5 Graphene supercell. (b) PDOS of O_2/Graphene system.

Finally, we consider the –OH radical. The O-H bond creates a deep-lying filled σ state and a high-lying empty σ* state. The other broken O bond makes the unpaired electron of the radical. As O is very electronegative, this p state lies well below E_F of isolated graphene. More interestingly, this p state is able to form a strong C-O bond to a carbon atom underneath, puckering the C atom out of the plane, and converting it into an sp^3 configuration (Fig 5-10). Thus, there is charge transfer from the graphene. However, the overall effect on conductivity will be poor because the defect states will lower the mobility.

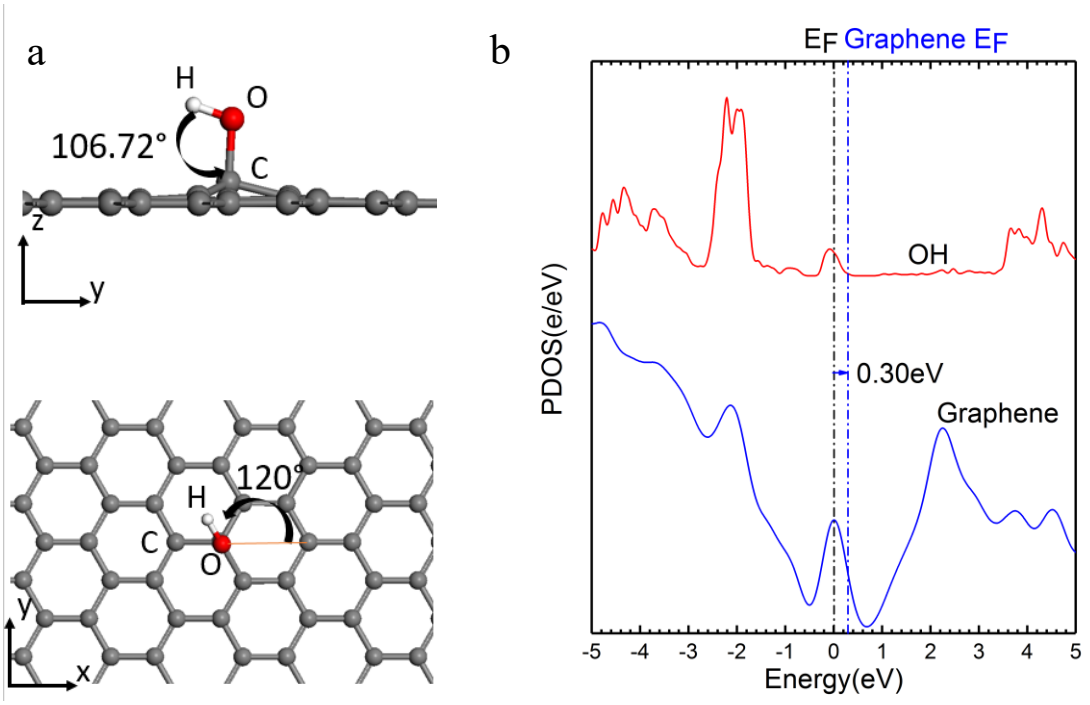


Fig 5-10. (a)Side view and (b) top view of OH radical bonding onto 5x5 Graphene supercell. (c) PDOS of OH/Graphene system.

Overall, except for OH, the various dopants studied are physisorbed, without puckering the underlying graphene. This occurs because of the strong intra-layer rigidity of graphene, and its resistance to out-of-plane deformation needed to form the fourth extra bond to a chemisorbing species.

5.4 Discussion

The electron affinity and ionization potentials of the various dopant species were calculated using dopant supercells as described in Section 5.2. The Fermi level shifts (FLS) are compared with the ionization potentials (IP) in Table 5-3. The SbF_5, $FeCl_3$ and $AuCl_3$ species have remarkably large ionization potentials if the band gaps are added to the work functions.

We see that there is a monotonic variation of the calculated FLS with the IP. The largest calculated p-type shift occurs for SbF_5 and $FeCl_3$ has the largest shift of the more common

dopants $FeCl_3$, $AuCl_3$, and HNO_3. Experimentally, $FeCl_3$ is found to give the largest E_F shift of the common dopants $FeCl_3$, $AuCl_3$, MoO_3, and HNO_3 [38, 39].

For MoO_3 doping, our calculations suggest there is no puckering of the underlying C site, so there will be no Raman D peak, and no extra carrier scattering. This is consistent with experiment where notably Chen *et al.* [17] find that MoO_3 doped graphene retains the ability to show quantum Hall effect, indicating a high carrier mobility.

For $FeCl_3$ doping, our calculations suggest there is no C site puckering, so there will be no Raman D peak and no extra carrier scattering. This is consistent with experiment [28, 37-43] although a small D peak does appear in some cases [41, 42].

Also, the absence of a Raman D peak at 1350 cm^{-1} in the experiment works for $FeCl_3$ [28], confirms that $FeCl_3$, MoO_3, and HNO_3 do not give rise to basal plane defects [17, 20, 37], and thus should not increase carrier scattering.

Our calculations have a similar aim as those of Hu and Gerber [37]. For $FeCl_3$, our calculations are for the expected spin-polarized state using GGA+U whereas Liu *et al.* [39] used a spin unpolarised state. We used a more efficient, three times smaller supercell than Zhan *et al.* [37] did by rotating the x,y-axes. Overall, the shift of E_F seen in the various calculations of $FeCl_3$ is similar. For HNO_3 doping, we found that the acid dissociates, as in Nistor *et al.* [47]. This work has considered the widest range of dopant species, including n-type dopants, compared them, and also studied the C site puckering, because it is no use increasing carrier density by doping, if the mobility declines by a similar factor. The main factor that leads to puckering is that the bond to carbon is too strong, for example, from an oxygen radical, and is to be avoided for the most effective form of doping.

5.5 Conclusion

We have calculated the conditions required for charge transfer doping of graphene (sometimes called non-covalent doping). We find that the Fermi level shift in eV is proportional to the electron affinity of the acceptor species or ionization potential of the donor species. We have treated a wider range of dopant species that other groups. Except for the case of –OH radicals, the dopants physisorb onto the graphene and thus do not create sp^3 “defects” and do

not degrade the mobility or cause Raman D peaks. The doping mechanism is similar to that occurring in transfer doping of organic semiconductors.

REFERENCES

1. A. K. Geim and K. S. Novoselov, Nat. Mater. **6** 183 (2007).
2. S. De and J. N. Coleman, ACS Nano **4** 2713 (2010).
3. S. Bae, H. Kim, and Y. Lee, Nat. Nanotechnol. **5** 574 (2010).
4. G. Zhong, X. Wu, L. D'Arsie, and J. Robertson, Appl. Phys. Lett. **109** 193103 (2016).
5. S. Sato, Jpn. J. Appl. Phys. Part 1 **54** 040102 (2015).
6. S. Hofmann, P. B. Weimer, and R. S. Weatherup, J. Phys. Chem. Lett. **6** 2714 (2015).
7. F. Schedin, A. K. Geim, S. V. Morozov, E. W. Hill, P. Blake, M. I. Katsnelson, and K. S. Novoselov, Nat. Mater. **6** 652 (2007).
8. T. O. Wehling, K. S. Novoselov, S. V. Morozov, E. E. Vdovin, M. I. Katsnelson, and A. K. Geim, Nano Lett. **8** 173 (2008).
9. R. T. Lv, G. G. Chen, Q. Li, A. McCreary, A. Botello-Mendez, S. V. Morozov, L. B. Liang, X. Declerk, N. Perea-Lopez, D. A. Culleni, K. S. Novoselov, and M. Terrones, Proc. Natl. Acad. Sci. U.S.A. **112** 14527 (2015).
10. S. Kasap, Principles of Electronic Materials and Devices (McGraw-Hill, NY, 2000), p. 351.
11. Y. C. Lin, C. Y. Lin, and P. W. Chiu, Appl. Phys. Lett. **96** 133110 (2010).
12. D. Wei, Y. Liu, Y. Wang, H. Zhang, L. Huang, and G. Yu, Nano Lett. **9** 1752 (2009).
13. R. Lv, A. R. Botello-Mendez, T. Hayashi, B. Wang, A. Berkdemir, Q. Hao, A. L. Elias, R. Cruz-Silva, H. R. Gutierrez, Y. A. Kim, H. Muramatsu, H. Terrones, J. V. Charlier, and M. Terrones, Sci. Rep. **2** 586 (2012).
14. Y. C. Lin, P. Y. Teng, C. H. Yeh, M. Koshino, P. W. Chiu, and K. Suenaga, Nano Lett. **15** 7408 (2015).
15. T. Sharifi, G. Hu, X. Jia, and T. Wagberg, ACS Nano **6** 8904 (2012).
16. A. G. Rinzler, J. Liu, H. Dai, P. Nikolaev, C. B. Huffman, F. J. Rodriguez-Macias, P. J. Boul, A. H. Lu, D. Heymann, and D. T. Colbert, Appl. Phys. A **67** 29 (1998).
17. Z. Chen, I. Santodo, R. Wang, L. F. Xie, H. Y. Mao, H. Huang, Y. Z. Wang, X. Y. Gao, Z. K. Chen, A. T. S. Wee, and W. Chen, Appl. Phys. Lett. **96** 213104 (2010).
18. J. Meyer, P. R. Kidambi, B. C. Bayer, C. Weijtens, A. Kahn, A. Centeno, A. Pasquera, Z. Zurutuza, J. Robertson, and S. Hofmann, Sci. Rep. **4** 5380 (2014).
19. S. Sanders, A. Cabrero-Vilatela, P. R. Kidambi, J. R. Alexander-Webber, C. Weijtens, P. Braeuninger-Weimer, A. L. Aria, M. Oasim, T. D. Wilkinson, J. Robertson, S. Hofmann, and J. Meyer, Nanoscale **7** 13135 (2015).
20. L. D'Arsie, S. Esconjauregui, R. S. Weatherup, X. Wu, W. E. Arter, H. Sugime, C. Cepek, and J. Robertson, RSC Adv. **6** 113185 (2016).
21. S. Esconjauregui, L. D'Arsie, Y. Guo, J. W. Yang, H. Sugime, S. Caneva, C. Cepek, and J. Robertson, ACS Nano **9** 10422 (2015).
22. S. L. Hellstrom, M. Vogueritchian, R. M. Stoltenberg, I. Irfan, M. Hammock, Y. B. Wang, C. C. Jia, X. Guo, Y. L. Gao, and Z. Bao, Nano Lett. **12** 3574 (2012).
23. D. Kondo, H. Nakano, B. Zhou, I. Akiko, K. Hayashi, M. Takahashi, S. Sato, and N. Yokoyama, in Proceedings of the IEEE International Interconnect Technology Conference (IITC), 2014, p. 189.
24. U. Detlaff-Weglikowski, V. Skakalova, R. Graupner, S. H. Jhang, B. H. Kim, H. J. Lee, L. Ley, J. W. Park, D. Tomanek, and S. Roth, J. Am. Chem. Soc. **127** 5125 (2005).
25. D Kiriya, M Tosun, P Zhao, J S Kang and A Javey, J. Am. Chem. Soc. **136** 7853 (2014).

26. I Salzmann, G Geimel, M Oehzelt, SWinkler and N Koch, Acc. Chem. Res. **49** 370 (2016).
27. M. S. Dresselhaus and G. Dresselhaus, Adv. Phys. **30** 139 (1981).
28. W. Zhao, P. H. Tan, J. Liu, and A. C. Ferrari, J. Am. Chem. Soc. **133** 5941 (2011).
29. J. Meyer, R. Khalandovsky, P. Görrn, and A. Kahn, Adv. Mater. **23** 70 (2011).
30. J. Meyer, S. Hamwi, M. Kroger, W. Kowalsky, T. Reidl, and A. Kahn, Adv. Mater. **24** 5408 (2012).
31. S. J. Clark, M. D. Segall, and C. J. Pickard, Z. Kristallogr. **220** 567 (2005).
32. S. J. Clark and J. Robertson, Phys. Rev. B **82** 085208 (2010).
33. A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102** 073005 (2009).
34. S. Grimme, J. Comput. Chem. **27** 1787 (2006).
35. T. Hu and I. C. Gerber, J. Phys. Chem. C **117** 2411 (2013).
36. W. E. Arter, L. D'Arsie, X. Wu, C. S. Esconjauregui, and J. Robertson, Appl. Phys. Lett. **110** 223104 (2017).
37. D. Zhan, L. Sun, Z. H. Ni, L. Liu, X. F. Fan, Y. Wang, T. Yu, Y. M. Lam, W. Huang, and Z. X. Shen, Adv. Funct. Mater. **20** 3504 (2010).
38. Y. Song, W. Fang, A. L. Hsu, and J. Kong, Nanotechnology **25** 395701 (2014).
39. W. Liu, J. Wang, and K. Banerjee, IEEE Electron Device Lett. **37** 1246 (2016).
40. K. K. Kim, A. Reina, Y. Shi, H. Park, L. J. Li, Y. H. Lee, and J. Kong, Nanotechnology **21** 285205 (2010).
41. K. C. Kwon, K. S. Choi, and S. Y. Kim, Adv. Funct. Mater. **22** 4724 (2012).
42. X. Meng, S. Tongay, J. Kang, Z. Chen, F. Wu, S. Li, J. B. Xia, J. Li, and J. Wu, "FeCl3 doping," Carbon **57** 507 (2013).
43. Y. Li and Q. Yue, Physica B **425** 72 (2013).
44. S. Chuang, C. Battaglia, A. Azcatl, S. McDonnell, K. S. Kang, R. M. Wallace, and A. Javey, Nano Lett. **14** 1337 (2014).
45. Y. Guo, D. Liu, and J. Robertson, ACS Appl. Mater. Inter. **7** 25709 (2015).
46. Y. Guo and J. Robertson, Appl. Phys. Lett. **105** 222110 (2014).
47. R. A. Nistor, D. M. Newns, and G. J. Martyna, ACS Nano **5** 3096 (2011).

Chapter 6 Band edge states, intrinsic defects and dopants in monolayer HfS_2 and SnS_2

In this chapter, we focus on electronic properties, defect formation and dopant of 2D HfS_2 and SnS_2. Although monolayer HfS_2 and SnS_2 do not have a direct band gap like MoS_2, they have much higher carrier mobilities. Their band offsets are favorable for their use with WSe_2 in tunnel field effect transistors (TFET). Here, we study the effective masses, intrinsic defects and substitutional dopants of these dichalcogenides. We find that HfS_2 has surprisingly small effective masses for a compound that might appear partly ionic. The S vacancy in HfS_2 is found to be a shallow donor while that in SnS_2 is a deep donor. Substitutional dopants at the S site are found to be shallow. This contrasts with MoS_2 where donors and acceptors are not always shallow or with black phosphorus where dopants can reconstruct into non-doping configurations. It is pointed out that HfS_2 is a more favorable than MoS_2 for semiconductor processing because it has more convenient CVD precursors previously used to make HfO_2.

6.1 Background

Open-shell transition metal dichalcogenides (TMDs) such as MoS_2 have been intensively researched as important two-dimensional semiconductors. Their band gap changes from an indirect gap for the bulk to direct gap for the monolayer case [1]. The relatively small dielectric screening in the monolayer case means that complex exciton behaviour becomes important even for relatively small band gaps [2]. Carriers can be manipulated between the degenerate valley states in valleytronics [3]. TMDs are also interesting as photo- and molecular sensors. On the other hand, for purely electronic devices, we can consider other layered semiconductors which might have a carrier higher mobility.

One proposed electronic device is the heterojunction tunnel field effect transistor (TFET). In this case, the continued scaling of transistors for computation creates a need for very low power

switches, in particular switches with a steep subthreshold slope below the thermionic limit of 60 mV/decade of a normal field effect transistor [4-7]. We note that TFETs operating in the subthreshold regime are also very sensitive sensor amplifiers [8]. TFETs would normally be built using heterojunctions of two lattice-matched III-V semiconductors with a staggered or broken-gap band alignment. However, the lattice-matching condition is not always met and this leads to interfacial mismatch defects which degrade switching performance. An alternative is to use stacked layer heterojunctions of two TMDs. TMDs offer a wide range of band gaps and band offsets [9-12], and due to their van der Waals inter-layer bonding, no lattice matching condition is needed to avoid dangling bond-type defects. Considering the band offsets of the various 2D semiconductors, a suitable choice is a p-type compound with d^2 configuration such as WSe_2 paired with an n-type d^0 compound such as HfS_2 or SnS_2 [9-12]. WSe_2 has a suitable high ionization potential whereas HfS_2 and SnS_2 have suitably deep electron affinities. SnS_2 and HfS_2 have s,p-like band edges and so, like black phosphorus, they have a higher phonon limited mobility of order 1000 $cm^2/V.s$ and lower effective masses than MoS_2 [13, 14].

The electronic properties of HfS_2 or SnS_2 are much less studied than the standard TMDs such as MoS_2 or WSe_2, It is particularly important to understand the intrinsic defects and anion vacancies of these materials because these defects can cause Fermi level pinning at the contacts [15-19], which causes a large contact resistance. This is a principal cause of the under-performance of 2D devices [20, 21]. Thus, this chapter investigates the band edge states, the intrinsic defects and the substitutional dopants of these chalcogenides.

6.2 Methods

The calculations are carried out with the CASTEP plane-wave density functional theory (DFT) code [22, 23] for periodic supercell models of the Hf/Sn disulphides. Ultra-soft pseudopotentials are used and the Perdew-Burke-Ernzerhof (PBE) form of the generalized gradient approximation (GGA) is used for the electronic exchange-correlation functional for geometry relaxation. The HSE (Heyd-Scuseria-Ernzerhof) [24] hybrid functional is used to calculate the band structures and the heats of formation. The HSE parameters α and ω are set to 0.2 as in HSE06 to give band gaps consistent with experimental values. The screened

exchange (sX) method [23] is also used for heats of formation, for comparison. Spin-orbital coupling is not included. The plane wave cut-off energy is set as 260 eV. All atomic structures are relaxed to a residual force of less than 10^{-5} eV/atom. Van der Waals corrections [25, 26] are included for bulk structures.

For the 2D Hf/Sn disulphide system, a convergence test finds that a vacuum layer thickness of 20Å in the z-direction is enough to converge the formation energy of S vacancy and that 5x5 supercells in the x and y-direction are enough to allow us to neglect periodic images. The transition states of intrinsic defects are corrected using the Lany and Zunger scheme, illustrated in chapter 2 [27]. The formation energy of each charge state is given by

$$H_q(E_F,\mu) = [E_q - E_H] + q(E_V + \Delta E_F) + \sum_\alpha n_\alpha(\mu_\alpha^0 + \Delta\mu_\alpha) \tag{6.1}$$

where q is the charge on the system, E_q is the energy of the charged system with a defect, E_H is the energy of the charged defect-free system. E_V is the valence band maximum (VBM) and E_F is the Fermi level with the respect to VBM. n_α is the number of atoms of species α, and μ_α is the relative chemical potential of element α. We note that the first two terms are equal to the difference between the total energy of charged defect system and total energy of the neutral defect-free system.

6.3 Results

6.3.1 Band structure, alignment and effective mass

Each of HfS_2 and SnS_2 has the 2H structure with an octahedral metal site. The lattice constant of HfS_2 is calculated to be 3.68Å in PBE, which is 1.4% more than the experimental value of 3.62Å [28]. The lattice constant of SnS_2 is calculated to be 3.74Å, which is 2.9% more than experimental value of 3.64Å [28].

We then calculate the chemical potential for the S-rich and S-poor limits. In the S-rich limit, the chemical potential of S is set to 0 eV. In the S-poor limit for HfS_2, the S chemical potential is set to the Hf-HfS_2 equilibrium, from the heat of formation of HfS_2 (Table 6-1). This is calculated to be -5.10 eV or 2.55 eV/S atom in HSE, compared to -2.58 eV/S atom

experimentally [29]. For SnS_2, the monovalent sulphide SnS exists between SnS_2 and Sn metal [29-33], so the range of S chemical potential for SnS_2 is from 0 to -0.50 eV/ S atom, or 0 to -0.50 eV experimentally [31-33].

Table 6-1. Heats of formation by sX functional [29, 30].

	eV/mole
MoS_2	-3.04
HfS_2	-5.16
SnS_2	-1.53
SnS	-1.13

The band structures of bulk SnS_2 and HfS_2 have been studied for some time [34-38]. Fig 6-1 and 6-2 show the band structures of monolayer and bulk HfS_2 and SnS_2 calculated with the HSE functional. There is an indirect band gap in both monolayer and bulk forms, which is different from the case of MoS_2 and other d^2 transition metal dichalcogenides. The band gap is from Γ to M for the monolayer and from Γ to L for the bulk.

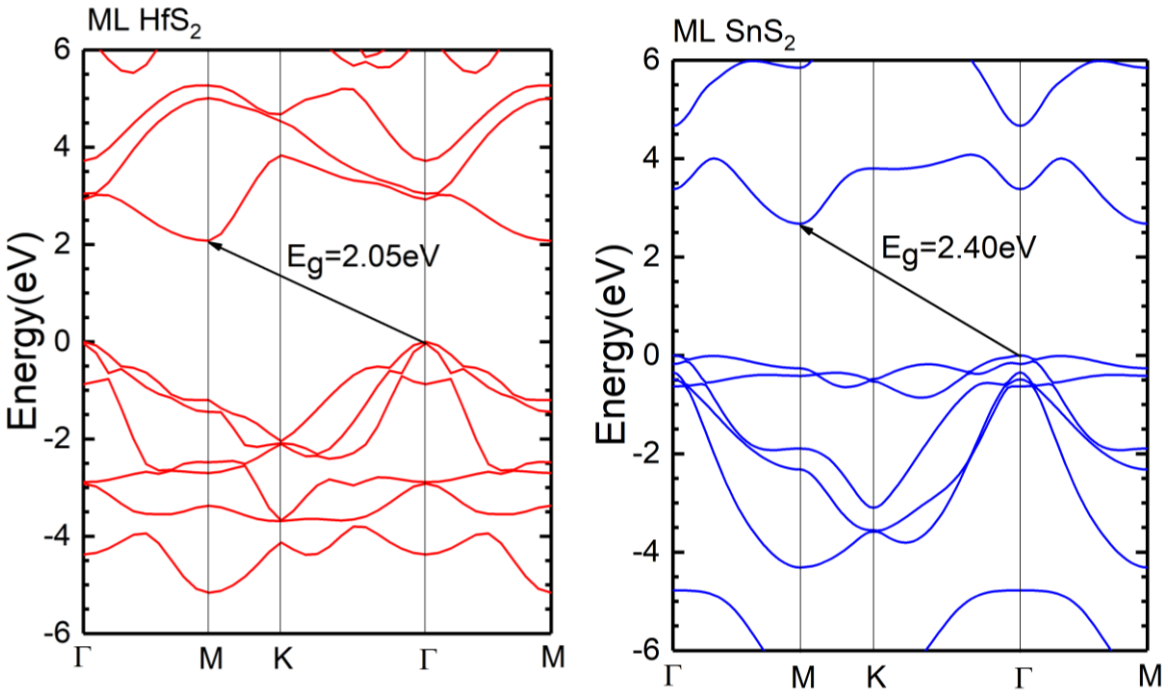


Fig 6-1. Band structures for the monolayers.

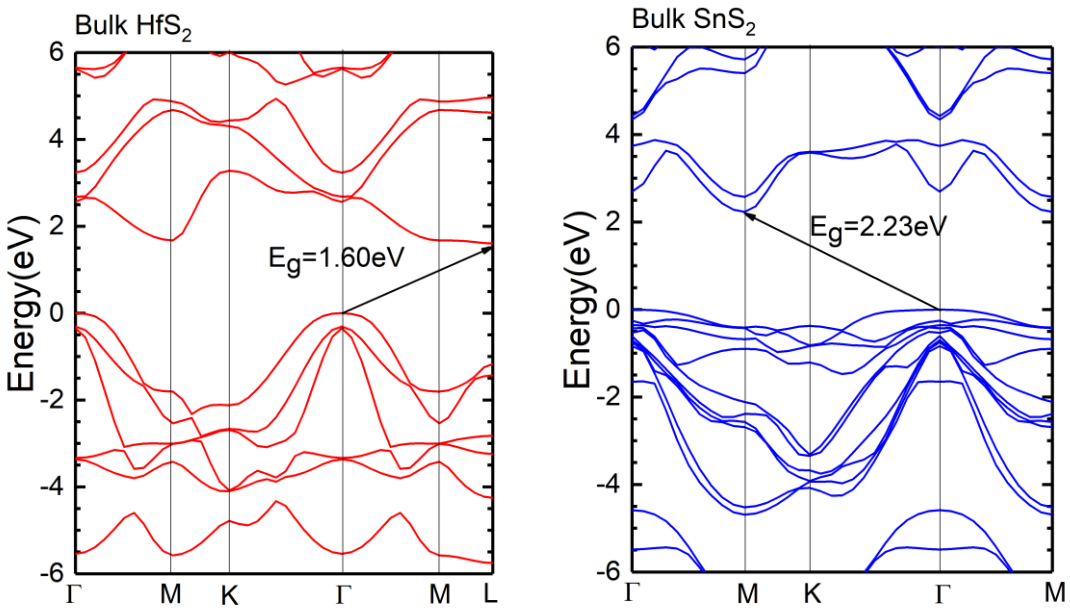


Fig 6-2. Band structures for the bulk compounds.

Table 6-2 compares the band gaps of these two materials calculated in PBE and HSE and the experimental band gaps for the bulk form [39]. Generally speaking, HSE06 corrects any under-estimation of the band gap of PBE.

Table 6-2 Calculated Band gaps of HfS_2 and SnS_2 compared to experimental values [12, 13, and 39]. ML = monolayer, CNL = charge neutrality level.

Band gaps (eV)	HfS_2		SnS_2	
	ML	bulk	ML	Bulk
PBE	0.98		1.50	
HSE	2.05	1.68	2.40	2.30
SX	2.12	1.95	2.68	2.0
Exp		1.98		2.18
CNL (ML)	1.11		1.55	

The calculated Bader charges are +0.34 for Hf in HfS_2 and +0.3 for Sn in SnS_2 showing that the bonding is relatively non-polar in these compounds despite the formal ionic charges often used to describe their bonding.

Table 6-3 shows the calculated effective masses for SnS_2 and HfS_2. The non-polar bonding (only 8% ionic for HfS_2) explains the relatively dispersed band structures and the small effective masses of these compounds. Our hole masses of SnS_2 differ slightly from those of Gonzalez [38].

Table 6-3 Effective masses for monolayer. e=electron, h=hole, x and y are along ΓK and ΓM respectively.

	HfS_2	SnS_2
m_ex	0.25	0.27
m_ey	1.85	0.72
m_hx	0.48	1.2
m_hy	0.49	2.8

Fig 6-3 shows the calculated band alignments with respect to the vacuum level [12]. These were calculated using supercells containing a monolayer of sulfide and 20Å of vacuum. This shows that WSe_2 has a type II band alignment with monolayer HfS_2 and SnS_2 in HSE as desired for a vertically stacked heterojunction TFET.

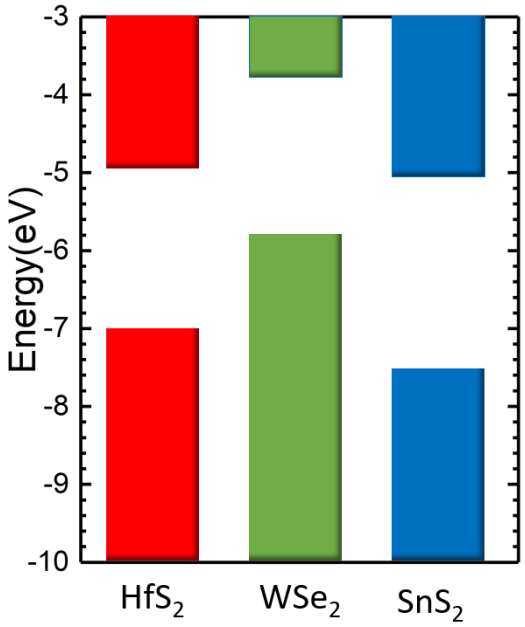


Fig 6-3. Calculated band offsets for stacked monolayers of HfS_2, SnS_2 and WSe_2.

6.3.2 Intrinsic defects

We now consider the geometries and formation energies of the intrinsic defects. Fig 6-4(a) shows the vacancy configuration. When the S atom is removed, the Hf or Sn and S atoms around the vacancy all move slightly away from vacancy centre, compared to the defect-free configuration. Fig 6-4(b) shows the formation energy as a function of Fermi energy E_F in the S-poor limit and the charge transition states. Here, the energies are plotted with respect to the charge neutrality level (CNL) [40] to enable both compounds to be plotted in a single diagram. For HfS_2, the -2 state is stable across all of the gap and with no state in the gap. The transition state lies at the bottom of the conduction band, so the vacancy is a shallow donor. For SnS_2, there is a transition level for –2 to +2 in the upper gap at +0.3 eV above the CNL. This vacancy is a deep donor.

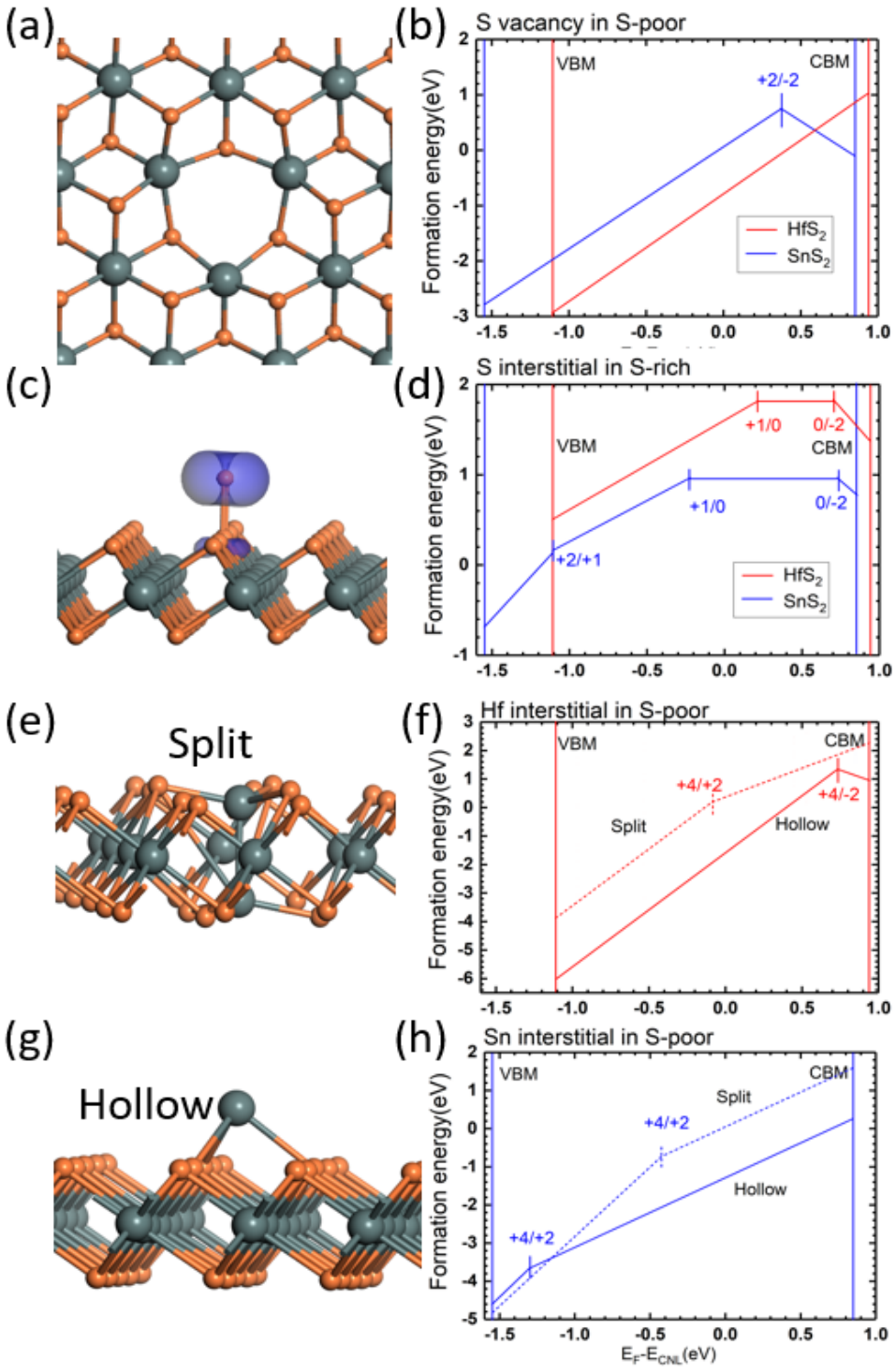


Fig 6-4. Geometries and defect formation energies vs Fermi energies for (a) and (b) S vacancy, (c) and (d) S adatom interstitial, (e) geometry split interstitial, (g) hollow interstitial for HfS_2 and SnS_2. (f) and (h) Compares the formation energies for the metal interstitial defect for (f) HfS_2 and (h) SnS_2 alone.

Fig 6-5(a) shows the partial density of states (PDOS) of the neutral defect state. For HfS_2, there is a peak in the PDOS at the conduction band edge with E_F lying at the conduction band edge. For SnS_2, transition state +2/-2 lies in the upper gap, above a defect band, consistent with Fig 6-4. The tendency to lose two electrons is the same for HfS_2 except for that Fig 6-5(b) now has two PDOS peaks for the +2/0 and 0/-2 states.

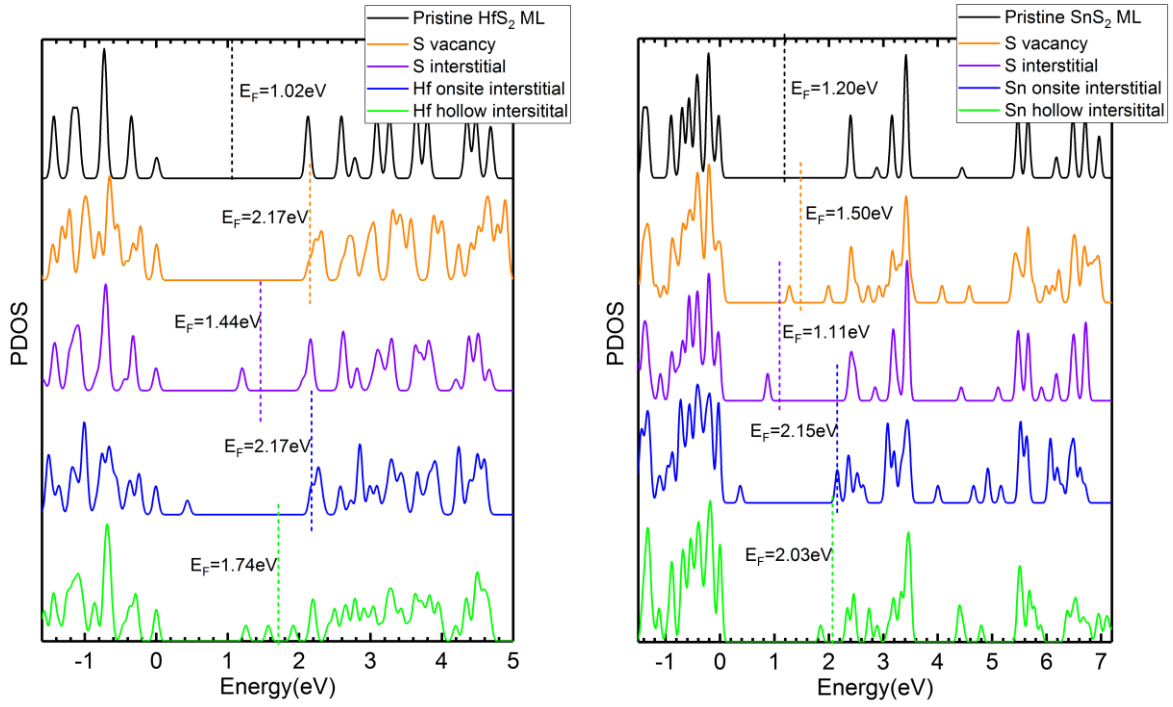


Fig 6-5. Partial density of states for the defect-free monolayer, S vacancy, S interstitial, metal split interstitial, and metal hollow interstitial, all in their neutral states, for HfS_2 and SnS_2.

The behaviour of the S vacancy in the d^0 compound HfS_2 differs from that of the S vacancy in the d^2 compounds MoS_2 where the neutral vacancy has a donor state in the upper band gap and a filled state at the valence band edge [18, 41].

The sulphur interstitial configuration is shown in Fig 6-4(c). This adatom configuration is found in many layered compounds. The S-S bond is calculated to be 1.99Å in HfS_2 and 1.98Å in SnS_2. The S-S bond is longer than the double bond and shorter than the S-S single bond in S_8. Fig 6-4(d) shows the formation energies and transition state of this defect in HSE06. PBE

gives three transition states in the gap +2/+1, +1/0, 0/-2, while HSE shows two defect states, +1/0, 0/-2. The +1/0 state lies in the middle of the gap, and the 0/-2 state lies at the conduction band edge. The orbitals for +2/+1 and +1/0 states are also shown in Fig 6-4(c). The +2/+1 orbitals consist of degenerate p_x and p_y states of the S adatom. The +1/0 orbitals have the same two degenerate orbitals but more located in underlying S atom. HSE gives a similar result, but only the +1/0 state is found, lying 0.7 eV below the conduction band minimum (CBM). This behaviour is similar to the S interstitial in monolayer MoS_2 [41]. (It should be noted that the S interlayer interstitial in bulk SnS_2 has a slightly different, where it tries to bond to both layers [32, 33].)

The Hf interstitial has two configurations in monolayer HfS_2, as seen in Figs 6-4(e) and (g). (The Sn interstitial in SnS_2 has similar behaviour.) One configuration has two Hf atoms stacked vertically on top of each other called the 'onsite' or 'split interstitial'. The other configuration places the extra Hf atom outside the layer at the hollow centre of three S atoms in the 'hollow interstitial'. Their formation energies are shown as a function of E_F in Fig 6-4(d).

For the split interstitial, the adjacent S atoms move away from defect centre to allow space for the extra metal atom. The two metal atoms are equivalent for the split interstitial. These atoms form in-plane bonds with the three adjacent S atoms. The system is symmetric in the z-direction. There are 4 valence electrons on Hf and Sn, two of which form three bonds with S. The other electron forms a Hf-Hf or Sn-Sn bond. There is one unpaired electron left, which can easily ionize. Hence, the +2 charge system dominates. The two electrons in Hf-Hf or Sn-Sn bond ionize if E_F moves across the transition energy. Both HfS_2 and SnS_2 have a similar mid-gap +4/+2 transition state. A mid-gap peak is seen at 0.4eV in Fig 6-4 (c) and (d), where the transition state is located.

The symmetry in the z-direction is lost for the 'hollow interstitial'. Three adjacent S atoms distort outward and out of the plane. There are two unpaired electrons in the extra Hf/Sn atom. Fig 6-4(h) shows the transition states. HfS_2 has a +4/+2 transition near the VBM and SnS_2 has nearly no transition state. Overall, plotting the formation energy of both interstitials across the band gap, the hollow site is the lowest for HfS_2 and the lowest for SnS_2 except very close to the valence band.

The metal vacancy states have also been calculated. Their formation energies for the neutral defects for the S-rich (metal-poor) limit are 4.38 eV and 5.31 eV for HfS_2 and SnS_2,

respectively. These formation energies are much higher than for the other defects. Therefore, we conclude that Hf and Sn vacancies are not very important.

We have also calculated the formation energies in PBE. While PBE underestimates band gap and the formation energy, it usually gives the right location of charge transition state with respect to the CNL. As Hf/SnS_2 is used for the n-type layer of the TFET, E_F will lie close to the CBM. Each of the S vacancy, interstitial and Hf/Sn interstitial has a positive formation energy near the CBM, which means that they will not form spontaneously.

6.3.3 Substitutional doping

Fig 6-6 shows the substitutional doping states at the S site. The Br donor is calculated to be a shallow state, with a transition state near the respective band edge. The As acceptor is deeper but still reasonably close to the VBM. This is very desirable if these compounds are to be used for a TFET. The fact that neither of the dopant sites reconstructs into a non-doping configuration explains why these sites are basically shallow, unlike the case of dopants in black phosphorus [42].

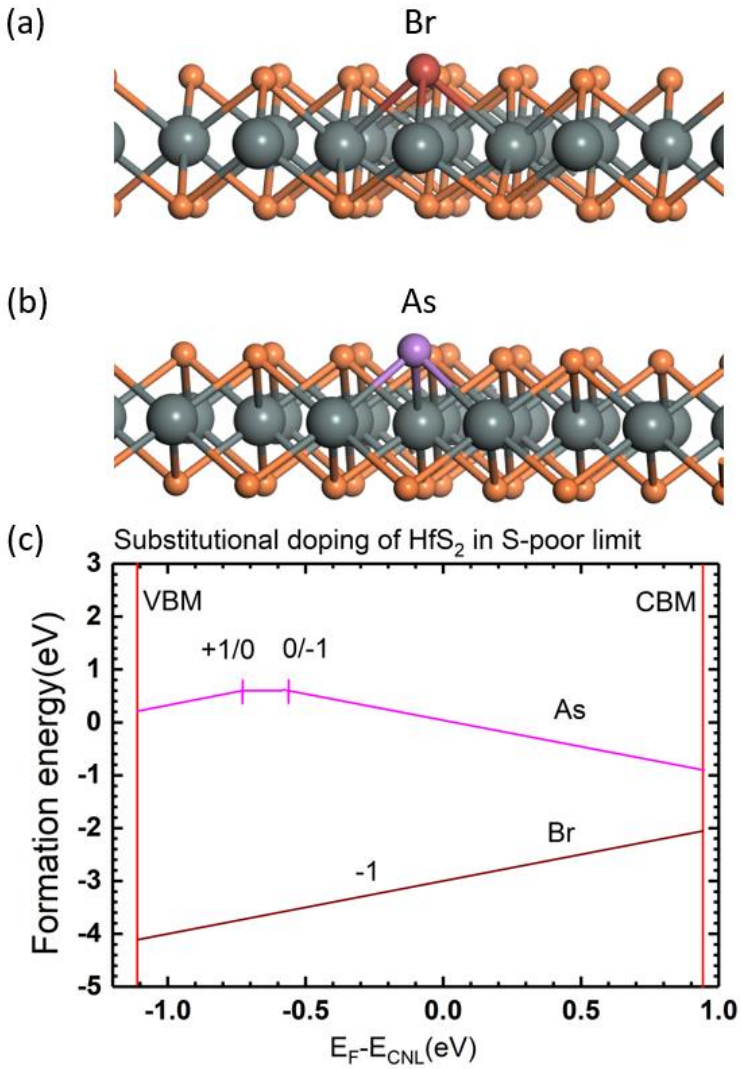


Fig 6-6. (a) and (b) Geometries of substitutional Br and As dopants at the S site in HfS_2, and (c) formation energy vs. Fermi energy.

6.4 Conclusion

We summarize the situation of these two compounds for use as a TFET. Their band offsets are as desired. SnS_2 has a low effective mass and is bipolar, with shallow donors and acceptors. Its main disadvantage is that it has only a small range of S chemical potential for which it is stable, which is important for growth by chemical vapour deposition (CVD). Superficially, HfS_2 is more ionic than SnS_2 and so, it might be expected to have higher effective masses.

However, in practice, its bonding is *not* very polar, and its effective masses are still low. Its big advantage is that it is the only stable sulphide of Hf, stable over a large stable range of S chemical potential. It has the great advantage that Hf CVD precursors are highly developed from the use of HfO_2 as a high K oxide in microelectronics, whereas precursors for MoS_2 like $Mo(CO)_6$ are less volatile and poisonous. The disadvantage of HfS_2 is that the S vacancy is a shallow donor. This will require CVD of HfS_2 to be carried out in S-rich conditions to increase the S vacancy formation energy and decrease its concentration. This might result in the formation of S interstitial adatoms, as already seen by Aretouli *et al*. [43]. Such adatoms may affect the quality of epitaxial growth. This would require careful control of S activity. Thus, HfS_2 is competing with InSe for use in TFETs. InSe has suitable band offsets, bipolar doping ability, and suitably behaved intrinsic defects [44], but maybe less convenient for CVD.

Finally, we have calculated the exfoliation energies for these compounds using the methods of Bjorkman *et al*. [45] and the Tkatchenko and Scheffler scheme [26] for van der Waals interactions. Our values in table 6-4 are similar to those found previously [45].

Table 6-4. Exfoliation energies (meV/Å^2)

HfS_2	SnS_2
22.2	10.2

In conclusion, HfS_2 and SnS_2 are indirect band gap semiconductors but otherwise very suitable for electronic devices because of their low effective masses and higher mobility than MoS_2. The high heat of formation makes it convenient for CVD. The main intrinsic defects in Hf/SnS_2 are the S vacancy, S interstitial and Hf/Sn interstitial. The S vacancy forms a gap state in SnS_2 and a shallow donor in HfS_2. The S interstitial is a low formation adatom. Substitutional dopants give reasonably shallow states. Therefore, both HfS_2 and SnS_2 can be considered as building blocks for TFETs.

REFERENCES

1. K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Phys. Rev. Lett. **105** 136805 (2010).
2. J. S. Rose, S. F. Wu, H. Y. Yu, N. J. Ghimiro, A. M. Jons, G. Aivarau, J. Q. Yan, D. G. Mandrus, D. Xiao, W. Yao, and X. D. Xu, Nat. Commun. **4** 1474 (2013).
3. K. F. Mai, K. He, J. Shan, and T. F. Heinz, Nat. Nanotechnol. **7** 494 (2012).
4. T. N. Theis and P. M. Solomon, Science **327** 1600 (2010).
5. D. Sarkar, X. J. Xie, W. Liu, W. Cao, J. H. Kang, Y. J. Gong, S. Kraemer, P. Ajayan, and K. Banerjee, Nature **526** 91 (2015).
6. T. Roy, M. Tosun, M. Hettick, G. H. Ahn, C. M. Hu, and A. Javey, Appl. Phys. Lett. **108** 083111 (2016).
7. S. Das, A. Prakash, R. Salazar, and J. Appenzeller, ACS Nano **8** 1681 (2014).
8. D. Sarkar, W. Liu, X. Xie, A. C. Anselmo, S. Mitragotri, and K. Banerjee, ACS Nano **8** 3992 (2014).
9. C. Gong, H. J. Zhang, W. H. Wang, L. Colombo, R. M. Wallace, and K. J. Cho, Appl. Phys. Lett. **103** 053513 (2013).
10. J. Kang, S. Tongay, J. Zhou, J. B. Li, and J. Q. Wu, Appl. Phys. Lett. **102** 012111 (2013).
11. Y. F. Liang, S. T. Huang, R. Soklaski, and L. Yang, Appl. Phys. Lett. **103** 042106 (2013).
12. Y. Guo and J. Robertson, Appl. Phys. Lett. **108** 233104 (2016).
13. Y. Huang, E. Sutter, J. T. Sadowski, M. Cotlet, O. L. A. Monti, D. A. Racke, M. R. Neupane, D. Wickramaratne, R. K. Lake, B. A. Parkinson, and P. Sutter, ACS Nano **8** 10743 (2014).
14. T. Kanazawa, T. Amemiya, A. Ishikawa, V. Upadhyaya, and Y. Miyamoto, Sci. Rep. **6** 22277 (2016).
15. S. McDonnell, R. Addou, C. Brule, R. M. Wallace, and C. L. Hinkle, ACS Nano **8** 2880 (2014).
16. C. M. Smyth, R. Addou, S. McDonnell, C. L. Hinkle, and R. M. Wallace, J. Phys. Chem. C **120** 14719 (2016).
17. Y. Guo, D. Liu, and J. Robertson, ACS Appl. Mater. Interfaces **7** 25709 (2015).
18. D. Liu, Y. Guo, L. Fang, and J. Robertson, Appl. Phys. Lett. **103** 183113 (2013).
19. Y. Guo, D. Liu, and J. Robertson, Appl. Phys. Lett. **106** 173106 (2015).
20. S. Das, H. Y. Chen, A. V. Penumatcha, and J. Appenzeller, Nano Lett. **13** 100 (2012).
21. A. Allain, J. H. Kang, K. Banerjee, and A. Kis, Nat. Mater. **14** 1195 (2015).
22. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, Z. Kristallogr. **220** 567 (2005).
23. S. J. Clark and J. Robertson, Phys. Rev. B **82** 085208 (2010).
24. J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118** 8207 (2003).
25. S. Grimme, J. Comput. Chem. **27** 1787 (2006).
26. A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102** 073005 (2009).
27. S. Lany and A. Zunger, Phys. Rev. B **78** 235104 (2008).
28. D. L. Greenaway and D. L. Nitsche, J. Phys. Chem. Solids **26** 1445 (1965).
29. H. Wada, K. Takada, and T. Sasaki, Solid State Ionics **172** 421 (2004).

30. O. Kubaschewski and C. B. Alcock, Metallurgical Thermochemistry (Pergamon, Oxford, 1979).
31. L. A. Burton, D. Colombara, R. D. Abellon, F. C. Grozema, L. M. Peter, T. J. Savenjie, G. Dennier, and A. L. Walsh, Chem. Mater. **25** 4908 (2013).
32. T. J. Whittles, L. A. Burton, J. M. Skelton, A. Walsh, T. D. Veal, and V. R. Dhanak, Chem. Mater. **28** 3718 (2016).
33. Y. Kumagai, L. A. Burton, A. Walsh, and F. Oba, Phys. Rev. Appl. 6, 014009 (2016).
34. C. Y. Fong and M. L. Cohen, Phys. Rev. B 5, 3095 (1972); J. Robertson, J. Phys. C **12** 4753 (1979).
35. L. F. Mattheiss, Phys. Rev. B **8** 3719 (1973).
36. H. Jiang, J. Chem. Phys. **134** 204705 (2011).
37. F. A. Rasmussen and K. S. Thygesen, J. Phys. Chem. C **119** 13169 (2015).
38. J. M. Gonzalez and I. I. Oleynik, Phys. Rev. B **94** 125443 (2016).
39. R. Schlaf, O. Lang, C. Pettenkofer, and W. Jaegermann, J. Appl. Phys. **85** 2732 (1999).
40. J. Robertson, J. Vac. Sci. Technol. B 18, 1785 (2000); Appl. Surf. Sci. **190** 2 (2002).
41. J. Y. Noh, H. Kim, and Y. S. Kim, Phys. Rev. B **89** 205417 (2014).
42. Y. Guo and J. Robertson, Sci. Rep. **5** 14165 (2015).
43. K. E. Aretouli, P. Tsipas, D. Tsoutsou, J. Marquez-Velasco, E. Xenogiannopoulou, S. A. Giamini, E. Vassalou, N. Kelaidis, and A. Dimoulas, Appl. Phys. Lett. **106** 143105 (2015).
44. Y. Guo and J. Robertson, Phys. Rev. Mater. **1** 044004 (2017).
45. T. Bjorkman, A. Gulans, A. V. Krasheninnkov, and R. M. Nieminen, Phys. Rev. Lett. **108** 235502 (2012).

Chapter 7 Passivation of the sulphur vacancy in monolayer MoS_2

Various methods to passivate the sulphur vacancy in 2D MoS_2 are modeled using density functional theory (DFT) to understand the passivation mechanism at an atomic scale. First, the organic super acid, bis(trifluoromethane)sulphonamide (TFSI) is a strong protonating agent and it is experimentally found to greatly increase the photoluminescence (PL) efficiency. DFT simulations find that the effectiveness of passivation depends critically on the charge state and number of hydrogens donated by TFSI, since this determines the symmetry of the defect complex. A symmetrical complex is formed by three hydrogen atoms bonding to the defect in a -1 charge state, and this gives no band gap states and a Fermi level in midgap. However, a charge state of +1 gives a lower symmetry complex with one state in the gap. One or two hydrogens also give complexes with gap states. Second, passivation by O_2 can provide partial passivation by forming a bridge bond across the S vacancy, but it leaves a defect state in the lower band gap. On the other hand, substitutional additions do not shift the vacancy states out of the gap.

7.1 Background

2D semiconductors such as the transition metal dichalcogenides (TMDs) have attracted considerable attention as optoelectronic devices because of their direct band gap in the monolayer form [1-3], and as alternatives to Si and III-Vs in field effect transistors because their thin layers allow excellent electrostatic control of their channels and so give good short channel performance [4, 5]. Their wide range of band gaps and band offsets give them potential for use as tunnel field effect transistors [6, 7]. Their interlayer van der Waals bonding means that the pristine systems in principle have no dangling bonds. However, a large concentration of defects (~10^{13} cm^{-2}), thought to be sulphur vacancies, is seen in transmission electron microscopy and scanning tunnelling microscopy (STM) on

exfoliated samples [8,9] and many like-atom bonds exist at the grain boundaries in samples grown by chemical vapour deposition (CVD) [10, 11]. Both types of defects will give rise to gap states and will reduce device performance. For example, sulphur vacancies are seen to reduce the photoluminescence efficiencies by typically 10^4 [12] while their field-effect mobility in devices is well below their phonon limited mobility [13] due to both high contact resistances [14] and defects.

In 3D semiconductors, there are strategies available to passivate defects. In MoS_2 ways to passivate defects have been tried with varying success, but there is presently no general understanding of how best to achieve this. The most successful passivation process so far has been treating the sample by an organic superacid bis(trifluoromethane) sulphonamide (TFSI) [12]. In this chapter, we study various possible passivation schemes for TMDs and explain why they are more complex than for simpler covalent semiconductors like Si.

Several passivation schemes for MoS_2 have been reported. (1) MoS_2 defect states can be removed by charge transfer doping via the van der Waals bonding of an organic monolayer [15]. (2) Thiol-based molecules can reduce the sulphur vacancy density on MoS_2 and achieve a high mobility of 80 $cm^2/(Vs)$ by a series of sulfurization reactions [16, 17]. Chemisorbed thiol groups can also achieve p-type or n-type doping by choosing different functional groups [18, 19]. (3) Molecular oxygen can passivate MoS_2 via chemisorption at the sulfur vacancy site [20]. This removes some vacancy gap states [21] and allows the photoluminescence (PL) efficiency to recover [22]. (4) Monolayer MoS_2 can be treated with organic super acid TFSI. This improves the PL quantum yield and efficiency [12, 23-25].

Passivation can be defined as a process removing all defect states from the gap while allowing the Fermi energy E_F to return to midgap. There are two standard methods to passivate defects in 3D semiconductors. (1) use a chemical reactant which bonds strongly with the defect so that the resulting electronic states now lie outside the gap [26-28]. (2) shift the defect states away from the relevant energy range [29-31]. Examples of the first method are the passivation of the residual Si dangling bonds at the Si/SiO_2 interface by hydrogen, where the resulting Si-H bonding and antibonding states lie in the valence band

(VB) and conduction band (CB) respectively [27]. Examples of the second method are adding InP capping layers to the active InGaAs channel layer, where the surface states of the InP layers lie outside the energy range of the InGaAs band gap while the original InGaAs gap states are now forming bulk bonds with states outside the gap [30, 31].

7.2 Methods

We now use a series of defect supercell calculations to investigate possible defect reactions. The atomic geometries and electronic properties of the three passivation schemes are calculated using the density functional theory (DFT) plane-wave CASTEP code [32, 33]. Ultra-soft pseudopotentials with a plane wave cut-off energy of 320eV are used. The Perdew-Burke-Ernzerhof (PBE) form of the generalized gradient approximation (GGA) is used as the electron exchange-correlation functional. The GGA treatment of the van der Waals interaction is corrected using the Grimme scheme [34]. Geometry relaxation is performed until the residue force is lower than 0.03eV/Å. A convergence test finds that a 4x4x1 supercell with a vacuum gap of 30Å and a 3x3x1 k-point mesh describe the 2D system with a single sulphur vacancy well. To overcome the error caused by the periodical mirror charge, a self-consistent dipole correction is implemented. Spin-polarization is used for molecular oxygen. Although molecular oxygen is a spin triplet, when it bonds onto the sulphur vacancy it becomes a singlet state.

The defect formation energies are calculated using the supercell method. Corrections for defect charges and band occupations are applied as in the Lany and Zunger scheme [35]. The total energy of the perfect host supercell (E_H) and the supercell with a defect (E_q) are calculated for different charge states. The defect formation energy H_q is then found from

$$H_q(E_F,\mu) = [E_q - E_H] + q(E_V + \Delta E_F) + \sum_\alpha n_\alpha(\mu_\alpha^0 + \Delta\mu_\alpha) \tag{7.1}$$

where q is the charge on the system and E_q is the energy of the charged system with a defect. E_H is the energy of the charged defect-free system, E_V is the valence band maximum

(VBM), and ΔE_F is the Fermi level with the respect to VBM. n_α is the number of species α, and μ_α is the relative chemical potential of element α.

The equilibrium lattice parameter of 2D MoS_2 is calculated in GGA to be 3.17Å, 0.7% error compared to experimental value [36]. The calculated GGA band gap is 1.72eV compared to an experimental optical band gap of 1.80 eV [1] and a calculated band gap of 1.88 eV in screened exchange (sX) [37]. Thus, GGA gives less band gap error for MoS_2 than for other layered chalcogenides like HfS_2 or InSe [38].

7.3 Results

7.3.1 Hydrogen passivation

For the defects, TFSI is known to greatly improve the PL efficiency [12]. It is dissolved in an organic alkane forming the TFSI anion and a nearly-free proton which can hop from anion to anion [39]. The TFSI anion is physisorbed near the S vacancy. It only forms a weak van der Waals bond, so it does not passivate directly. However, TFSI is a strong protonating agent with a large Hammett number. Its proton (H^+) is assumed to be the passivating agent. Fig 7-1 shows protons leaving the TFSI anion and approaching the S vacancy. The vacancy complex with protons can trap electrons if necessary to form a local closed-shell system.

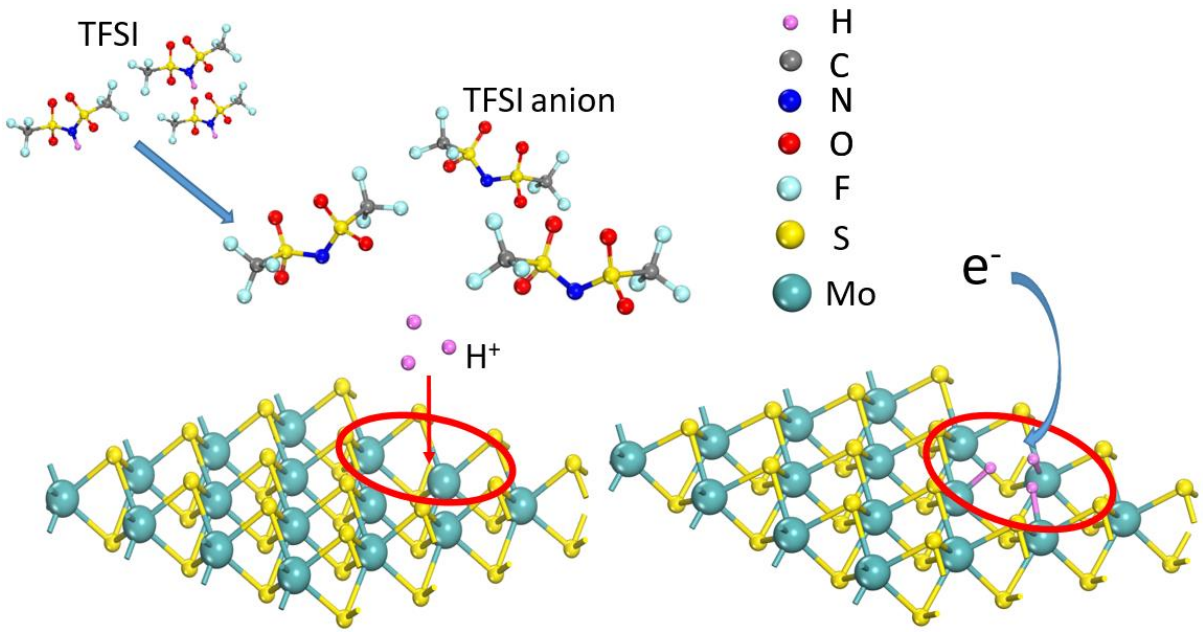


Fig 7-1. TFSI passivation schematic. The super-acid is a strong protonating agent where the protons can move freely to vacancy site and interact with Mo dangling bonds. In the Figure, three Hs are adsorbed onto the S vacancy where additional electrons can be trapped. The vacancy is at the center of the red circle.

To understand the adsorption configuration, proton passivation is modeled as a function of its charge state. Unlike in Si, the bonding in MoS_2 is multi-centred. One Mo dangling bond contributes only 2/3 of an electron to a Mo-S bond, rather than one electron as in a Si-Si bond. The vacancy site has trigonal C_{3v} symmetry, where three Mo dangling bonds form one resonant a_1 state and two degenerate e states around the gap, Fig 7-2(a) and (b) [37, 40].

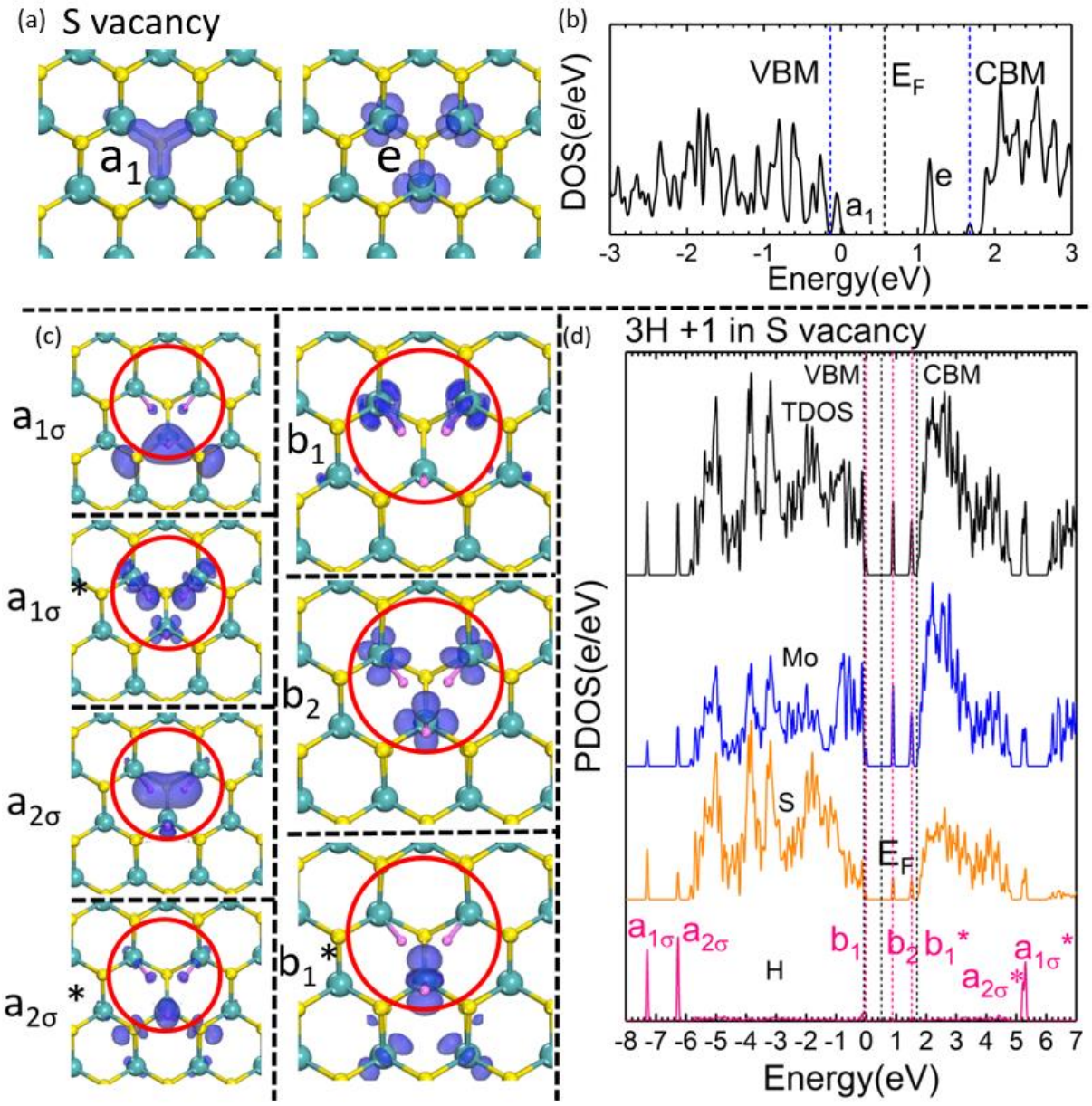


Fig 7-2. Simple S vacancy (top) and passivation by 3 hydrogens in the +1 state. (a) Defect orbitals and (b) density of states of the a1 and e gap states of the isolated S vacancy. (c) Asymmetric C_{2v} configuration for 3 protons at the vacancy with two electrons. Orbitals of the various localized states. The S vacancy lies at the center of the red circle. (d) PDOS showing the energies of the localized states.

To clean up the gap states, the symmetry should be conserved. Any half-filling of the e states breaks their spin degeneracy. Therefore, the adsorption configuration should be

closed-shell with trigonal symmetry, which needs three hydrogens. DFT modeling shows that the 3-H complex can either relax into the symmetrical or asymmetrical site, depending on the electron occupation of the complex. The +1 (2e) and -1 (4e) charged systems were considered by adding either 2 or 4 electrons to the system of the defect with three protons.

The +1 charged system is found to relax into an asymmetric C_{2v} configuration. Here, two hydrogens stay in the defect center while the third hydrogen moves away to an asymmetric off-centre site over a Mo atom, see Fig 7-2(c). The first two H's form a filled b1 bonding state with two of the Mo dangling bonds, Fig 7-2(d). Its empty anti-bonding partner b1* lies in the gap just below the CBM. The second e state forms an empty b2 state in midgap, localized mainly on the Mo dangling bonds.

The a_1 state of the vacancy interacts with the symmetric combination of the three hydrogens to form the a_1 and a_1* states at -6.3 eV and 5.5 eV respectively, well away from the band gap, Fig 7-2(d) and Fig 7-3(a). The asymmetric hydrogen interacts with two sulphur atoms to form the a_2 and a_2* states at -7.4 eV and 5.4 eV respectively, also well away from the band gap. The orbital character to the a2 state is also seen in the partial density of states (PDOS) in Fig 7-2(d). The asymmetric geometry of the +1 geometry is driven by the need to keep an empty b_2 state.

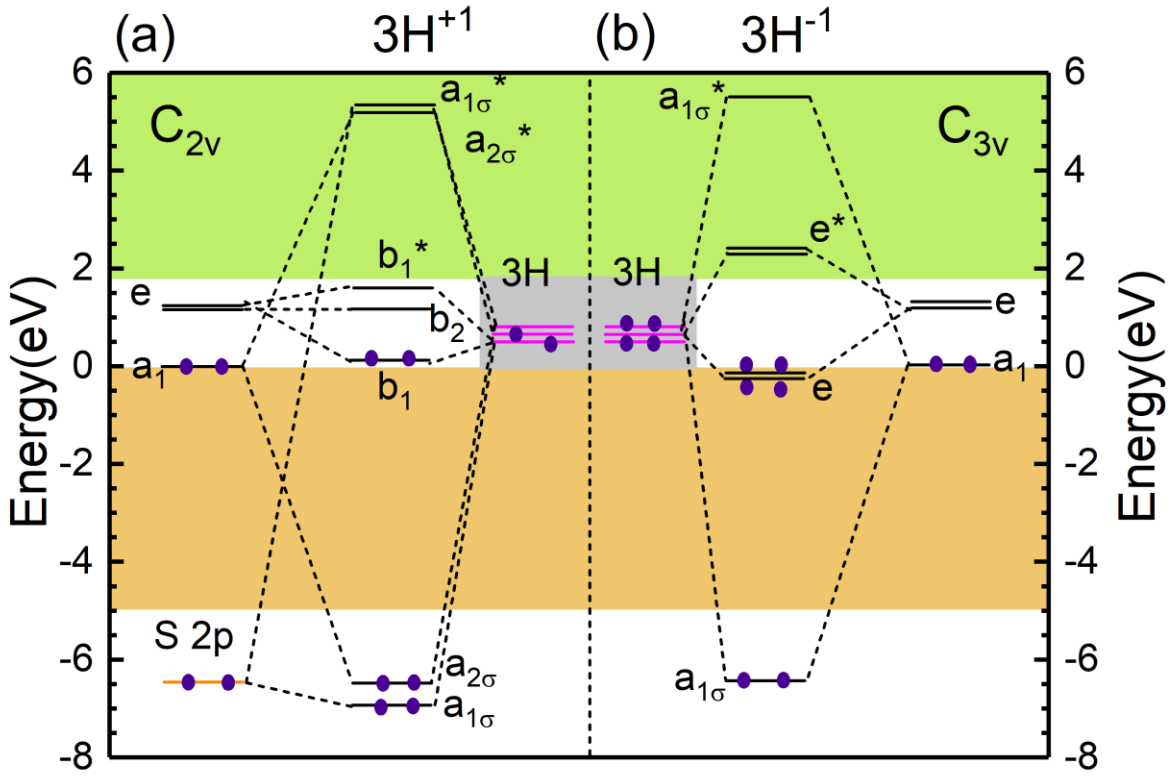


Fig 7-3. Molecular orbital diagram of 3 hydrogens interacting with S vacancy states. (a) Origin of a_1, a_2, b_1 and b_2 states for the C_{2v} 2 electron, +1 configuration. Some states remain in the gap. (b) Origin of the states of a_1 and e symmetry states for the C_{3v} 4-electron, -1 configuration. The energy levels of H atoms are in violet. The valence band is in orange and conduction band is green. All states are repelled from the gap in the -1 charge case because its higher symmetry causes an overall larger Mo-H interaction.

There is no passivation in the +1 state because of its lower symmetry. The off-centre hydrogen atom means that the $b_1 – b_1*$ splitting is too small to move these states out of the band gap, and the low symmetry and lack of interaction with hydrogens means that the b_2 state also remains in midgap.

To passivate all the e symmetry derived states, two more electrons should be added, giving a -1 charge (4e). Fig 7-4(a-f) shows the energetically favorable configuration, with three identical hydrogens and C_{3v} symmetry. The hydrogen orbitals form states of a_1 and e symmetry, which each interacts with Mo dangling bond orbital combinations of the same

symmetry, to form bonding states and anti-bonding states. The resulting a1 states lie deep in the valence band at -6.30 eV, and its anti-bonding partner a_1* lies well above the conduction band minimum (CBM) at 5.50 eV, Fig 7-3(d), Fig 7-4(b). Both of these orbitals extend along three local Mo-H bonds, Fig 7-4(a). The e states also form bonding and anti-bonding states, e and e*. The splitting of the e and e* states is now much larger than in the +1 case, and both states lie within the bands and outside the gap, as can be seen comparing Fig 7-3(a) and (b).

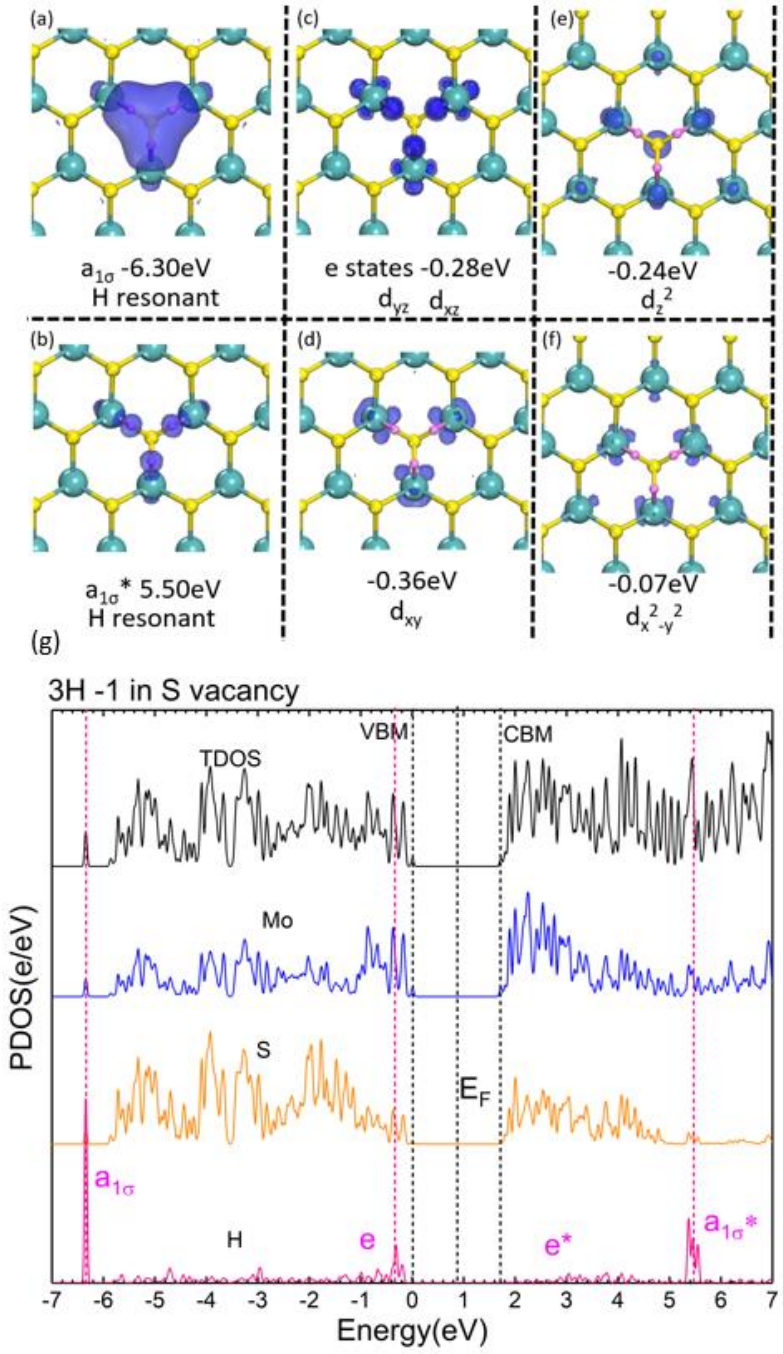


Fig 7-4. Passivation by 3 hydrogens in -1 charge state. (b) Orbitals of localized states in 3H - 1 passivated S vacancy with C3v symmetry: (a) hydrogen resonant bonding state a1σ which is below VBM, (b) hydrogen resonant anti-bonding state a1σ * located high in conduction bands, (c) doubly-degenerate bonding state of hydrogen with e states. Apart from hydrogen-related states, there are localized resonant states near the VBM (d) (e) (f), which are the d orbitals of the 3 adjacent Mo atoms. (g) PDOS showing various defect states.

Fig 7-4(g) shows the PDOS of the -1 state. This configuration repairs the S vacancy in MoS_2 and preserves the direct band gap of 1.72eV of perfect 2D MoS_2. The highest occupied molecular orbital (HOMO) is a delocalized Mo state, derived from pure MoS_2. The e state is a local resonant state. Both of them lie below the HOMO. Therefore, all gap states are removed and E_F lies at midgap, the 3H/4e passivation scheme is successful. The higher symmetry of this site has caused a larger bonding-antibonding splitting of the states to remove all the gap states.

Fig 7-5(a) shows the geometry for the vacancy with one hydrogen and a +1 charge state. For the 1H case, the hydrogen lies centrally in the vacancy, and this complex gives an empty state at midgap. For the 2H case in Fig 7-5(b), the hydrogens form Mo-H bonds with the dangling bonds and leave a single dangling bond with no hydrogen, and it gives a midgap state. Fig 7-5 also shows the formation energy vs E_F for each charge state, referenced to the H_2 chemical potential. This shows that the 1 H and 2H have a low formation energy. On the other hand, the 3H state has a slightly higher formation energy, but it is the only state which removes all gap states.

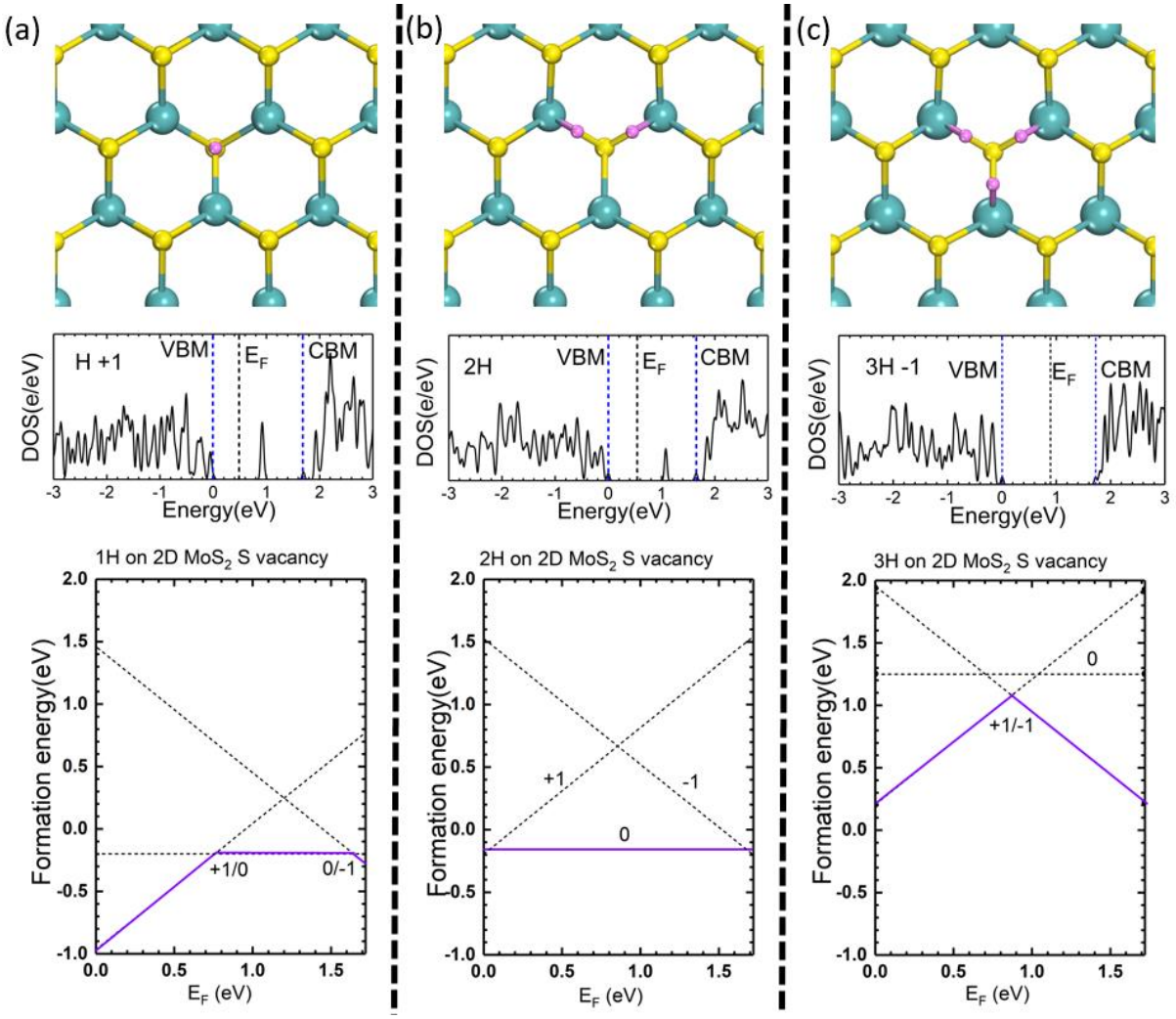


Fig 7-5. (a,b,c) Geometry, the local electronic density of states and defect formation energy diagrams of one, two and three hydrogens at the S vacancy in their H^+, $2H^0$ and $3H^-$configurations. The energies are referenced to the chemical potential of the H2 molecule.

The function of TFSI superacid is to supply a strongly acidic ambient to push the equilibrium toward greater binding of hydrogen with Mo dangling bonds. TFSI raises the chemical potential of H toward that of atomic H. This can be estimated from the Hammett acidity function of superacids as measured by Kutt *et. al* [50], as described in the appendix of this section, to roughly 0.5eV in Fig 7-5(c). This energy is above the stability line of the

-1 state so that electrons can be attracted to the defect if E_F is above the midgap (the +1/-1 transition state), due to background impurity levels often present in n-type MoS_2.

Hydrogen is not as effective a passivant of the S vacancy in MoS_2 as in Si-based systems because the energetics are less favorable, and that passivation occurs in only one charge state due to the complexity and symmetry effects associated with multi-centre bonding. The passivation efficiency is reduced by the ability of hydrogens to recombine into molecular hydrogen due to the ability of Mo-rich plane edge sites and the S vacancy sites to catalyze the hydrogen evolution reaction – which is favoured by only weak H binding energy [41, 42].

Hydrogen also binds to basal plane S sites, with the on-top site being the most stable. It is however 1.1eV less stable than at the S vacancy. The H can hop via the hollow site to an adjacent on-top site, with an energy barrier of only 0.1eV, as shown in Fig 7-6. In this way, protons denoted by TFSI are stable to diffuse and find S vacancy sites.

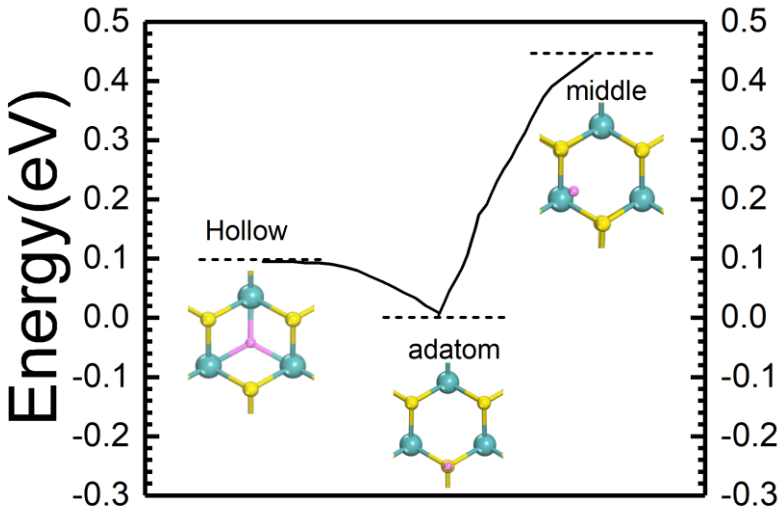


Fig 7-6. The hopping energy of hydrogen from one on-top S site to another. The zero-point is set as the energy of on-top S site.

While TFSI is a useful method to improve the optical properties of 2D MoS_2 samples, it also has weaknesses. The MoS_2 must be surrounded by a super-acid all the time to

maximize the hydrogen donation during device fabrication. Although a non-aqueous solvent or amorphous fluoropolymer binds the TFSI to the semiconductor [43], TFSI corrodes the electrode, even an inert metal.

7.3.2 Substitutional doping

We also consider other passivation methods. By analogy to passivation of the O vacancy in HfO_2, we can use two substitutional acceptors near the vacancy to compensate the loss of one S atom and to give the correct number of valence electrons to make a closed shell configuration, and return E_F to midgap again [44, 45]. This causes the vacancy to become V^{2+}. This charge causes a strong ionic relaxation around the vacancy which repels the vacancy state above the CBM, and so clears the gap of defect states. In MoS_2, the process involves either replacing two Mo atoms with two Nb atoms or replacing two adjacent S atoms with two As atoms, as in Fig 7-7(a) and (b). Fig 7-7(a) shows the atomic configuration and defect orbitals of Nb_{Mo} schemes. There are three different defect states, a_2: the asymmetrical resonant state, b_2: the d orbital of the local Mo hybrid with the Nb resonant state, e: the d orbital of the local Mo and Nb. Asymmetrical doping leads to the b_2 state and e state, shown in Fig 7-7(b).

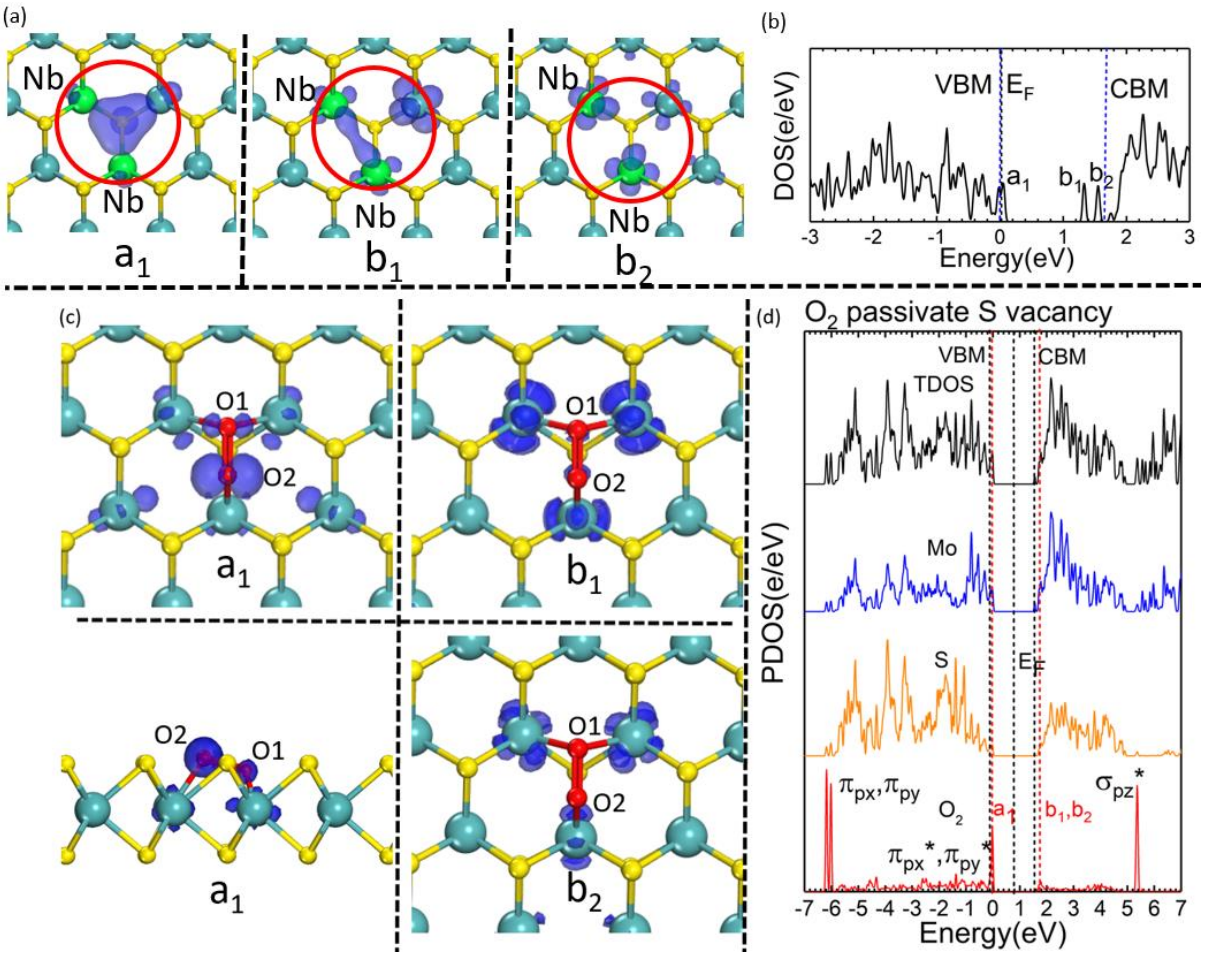


Fig 7-7. (a,b) Attempted passivation of S vacancy by two adjacent Nb/Mo sites, showing orbitals and PDOS. (c,d) Passivation of S vacancy by an O_2 molecule lying across the vacancy, orbitals and PDOS.

This scheme fails to passivate because the defect states do not move out of the gap. Despite its formal ionic charge of Mo^{+4}, the Mo-S bond in MoS_2 is not very ionic. The Bader charge of Mo in MoS_2 is actually only +0.22. Thus there is little ionic relaxation of Mo sites towards the positive vacancy as there was in HfO_2 to move the b_2 state out of the gap. Therefore, substitutional doping does not passivate the S vacancy.

7.3.3 Molecular passivation

7.3.3.1 Oxygen

The oxygen O_2 molecule is known to passivate the S vacancy experimentally. It is thought to occur by adding the undissociated O_2 molecule across the vacancy. The isolated neutral O_2 molecule is a spin triplet with two electrons with the same spin lying in the π_{px}* and π_{py}* states. Although the most stable configuration of molecular O_2 is open-shell, it becomes closed-shell when chemisorbed onto the S vacancy. As an undissociated molecule, one atom, O1, forms two Mo-O bonds and its second oxygen atom O2 forms one Mo-O bond. The three Mo-O bonds have the same length, 2Å, which allows O1 to lie inside the monolayer while atom O2 stays outside monolayer, as in the side view in Fig 7-7(c). This adsorption configuration is energetically stable after overcoming an energy barrier at room temperature [22]. The passivation occurs by compensating the S vacancy with the two unpaired π electrons from O_2. Bader charge analysis shows that charge is distributed evenly over the three Mo's, while O2 is more slightly electronegative than O1. Fig 7-7(c) shows the oxygen hybrid a1 and e states. Breaking the trigonal symmetry, the e states split up into b_1 and b_2 states, as shown in Fig 7-7(d). The a1 state is an oxygen state near the VBM, lying just in the gap. The b_1 and b_2 states are in the conduction band.

Despite the obvious advantage of the oxygen scheme which only needs neutral oxygen, the symmetry is broken, which means that the gap states are not sufficiently moved into the VB or CB. However, the a_1 state is fully occupied by four electrons as shown, so the vacancy is not a charge trap centre. The S vacancy may already be partially passivated during the growth of 2D MoS_2 since it is exposed to air. It is possible that increasing the oxygen density or raising the temperature to help O_2 over the adsorption energy barrier may achieve better passivation.

Atomic oxygen will also passivate the S vacancy, being 2.1eV more stable than adding S. Atomic O could be produced by a plasma or ozone. However, care would need to be taken that MoS_2 is not oxidized too far to a Mo oxide.

7.3.3.2 L-cysteine acid

It is natural to compensate the vacancy site with an extra sulphur interstitial, for example using a sulphur-containing molecule. One choice is a thiol, which has been reported to reduce the vacancy density thereby increasing the carrier mobility. The passivation involves either desulfurization [16] or dehydrogenation [18], which are complicated processes. Here, an alternative to L-Cysteine acid is considered, which only requires a reduction of L-cysteine. Cysteine acid is one of two sulfur-containing amino acids among 20 amino acids and the only amino acid with an –SH end group. It has two enantiomers, the levorotatory (L) and dextrorotatory (D) forms. Most cysteine acids found in nature are L-isomer, except in a few bacterial envelopes [46]. L-Cysteine plays an important part in biological electron transfer. Although L-form and D-form make no difference to the chemisorbed configuration, the enzyme can only be a catalyst of L-form reduction, which we use in this work: the dimer, L-cystine, can break its S-S bond whose energy is 1.77eV, with the help from cystine reductase, greatly reducing this chemical activation energy. This reaction is of great importance in cysteine metabolism [47]. The dimer then splits into two identical cysteine radical cations, as shown in Fig 7-8(a). It has been reported that the distonic-S form, which has an extra H^+ at NH_2-, is an energetically favorable isomer [48]. Then the cation adsorbs onto the S vacancy site, acting as S interstitial. Fig 7-8(a) illustrates these process and reaction pathway, whose final energy is 1.69eV lower than the initial configuration.

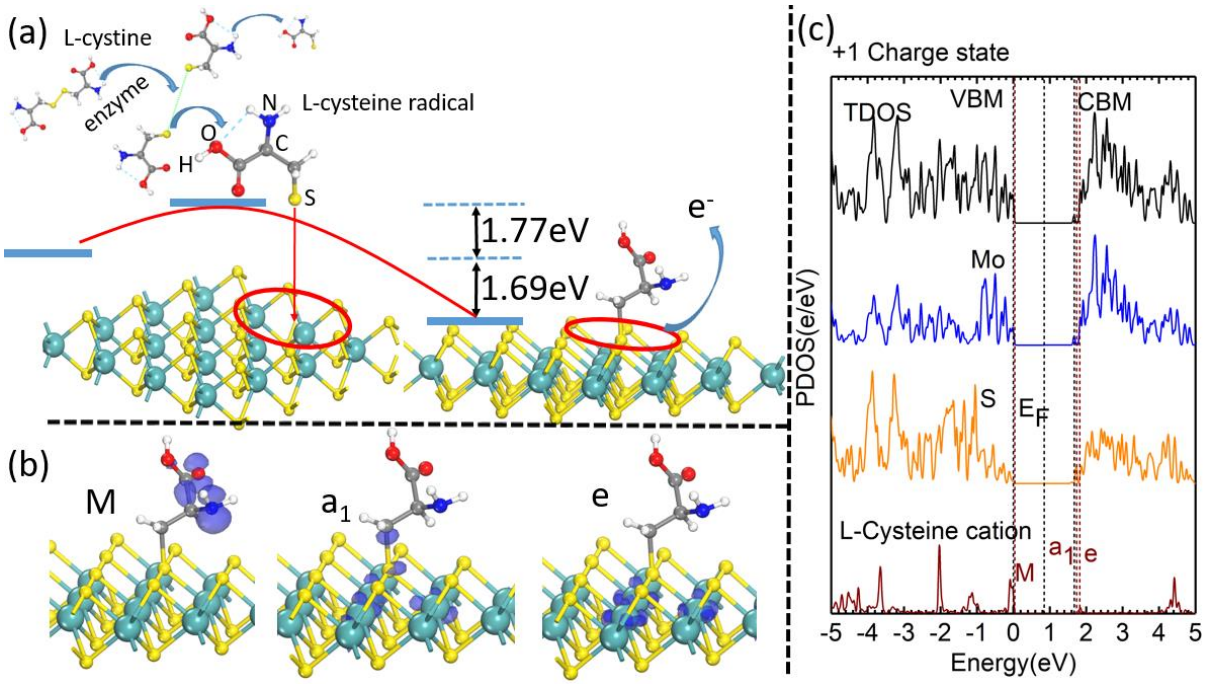


Fig 7-8. (a) L-Cystine breaks S-S bond into two dehydrogenated cysteine acids which approach the S vacancy of 2D MoS_2. (b) Orbitals of local states a_1, e, M out of the gap. (c) TDOS and PDOS of +1 state dehydrogenated (DH) cysteine acid adsorbed onto S vacancy.

As noted in the previous scheme, symmetry plays important role in whether the passivation is successful. Cysteine has no symmetry. However, only the S interstitial and β carbon affect the in-plane defect orbitals. The β carbon has two C-H and one C-C bond, which gives rise to slight symmetry breaking. Fig 7-8(b) shows the localized states near the band edge, which are molecular states M, the resonant state a_1 and d orbitals e. The S-C bond length is calculated to be 1.84Å, which is same as the length in cystine. Fig 7-8(c) illustrate that M state is at the VBM, occupied by four electrons, two of which come from the a_1 state of the S vacancy. These four electrons locate out of MoS_2 plane, therefore the M state has little effect on the property of MoS_2. Both the unoccupied states a_1 and e are symmetrical and locate above the CBM. According to Fig 7-8(b), a_1 is the resonant state of three local Mo, S interstitial and β carbon, and the shape of e states remains the same as in the S vacancy. Unlike hydrogen which cures

the defect states by bonding, cysteine acid mediates them by introducing a proper functional group with M states moving the defect states up to conduction band.

One advantage of this scheme is the maintaining of trigonal symmetry, which is one of the keys towards successful passivation. Apart from that, the S end in cysteine acid prefers to interact with the Mo dangling bonds rather than the d_z^2 bonds of Mo, which makes cysteine a crucial part in the self-assembled monolayer [49]. Although cysteine passivation involves tricky reaction steps and charged system, it is worthwhile investigating the dynamics of cysteine as surface cleaning molecular in the next work. Cysteine can potentially help move the S vacancy to the boundary by electrophoresis.

7.3.3.3 Others

The molecule titanyl phthalocyanine (TiOPC) can passivate the defect, where it causes a charge transfer to the defect which moves the vacancy orbital out of the gap into the bottom of the conduction band by an inter-molecular charge transfer, as has already been studied [15]. Finally, the thiol molecule also provides some passivation ability, by replacing the missing S atom of the vacancy. However, the efficiency of this process is not high [16].

7.4 Conclusion

In summary, we investigated the electronic properties of S vacancy and possible passivation schemes. Three hydrogen atoms symmetrically adsorbed around the S vacancy site in its –1 charge state successfully removes all gap states and returns E_F to midgap. The passivation mechanism is more complex than of covalently bonded systems like Si and SiO_2 because of the multi-centred bonding in MoS_2 and the resulting symmetry constraints that this imposes. Symmetry is critical to moving defect states out of the gap, to avoid lifting the defect state degeneracy, and because a sufficient energy splitting of bonding and anti-bonding states is needed to move states completely out of the gap. Other methods such as substitutional doping are not as effective because for example, the Mo-S bond is not as ionic as HfO_2.

7.5 Appendix: analysis of proton chemical potential due to TFSI

The Hammett acidity function (H_O) allows the calculation of the ratio of the real concentration of the base (B) and superacid (BH^+) in solution if you know the pK_a (called pK_{BH^+}). Normally all you can calculate from the pK_a are the activities. The activity (a) is the product of the mole fraction (x) and the activity coefficient (γ). In solutions, often $a_B=C_B\gamma_B$ is used, where C_B is the molarity. Once you know the H_O, you can calculate the proton activity if you know the activity coefficients of the acid and base. In a concentrated solution, the proton activity coefficients can be far from unity,

$$H_O = pK_{BH^+} + \log(\frac{[B]}{[BH^+]})$$

$$H_O = -\log(\frac{a_{H^+}\gamma_B}{\gamma_{BH^+}})$$

$$a_B = x_B \cdot \gamma_B$$

The chemical potential (μ) in an ideal solution is the chemical potential of the pure substance (μ^0) plus a correction proportional to the log of the mole fraction (x). However, in a real solution, the chemical potential is equal to the chemical potential of the pure substance plus a correction proportional to log(a). However, the activity is a functional of the concentration, (especially in a concentrated solution) so it must be measured. Therefore, there is no easy way to calculate the chemical potential of H_O,

$$\mu_B = \mu_B^0 + RT \cdot ln(x_B)$$

$$\mu_B = \mu_B^0 + RT \cdot ln(a_B)$$

To work around this, Kutt [50] developed a method to quantify protonation strength to the solvents (DCE and MeCN) and make measurements on 66 superacids. In solutions of DCE, the TFSI$(CF_3SO_2)_2NH$ (denoted as HA), will partially protonate DCE denoted as S (solvent). The pK_a value gives the relative amount of protonated solvent in the dilute solution.

$$HA + S \leftrightarrow A^- + SH^+$$

$$pK_a = -\log(\frac{a(SH^+) \cdot a(A^-)}{a(HA)})$$

For TFSI, K_a(DCE) is -11.9, while K_a(MeCN)=0.3; they are about 9.4 points lower than the values for H_2SO_4 [K_a=-2.5 (DCE) and 8.7 (MeCN)], a common reference. Therefore, although one cannot calculate the chemical potential of TFSI, one can say it is likely to raise the chemical potential of the proton donating species (SH^+) by $RT \cdot \ln(10^{-pKa})$ which at 25℃ equals to ~0.50eV.

REFERENCES

1. K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Phys. Rev. Lett. **105**, 136805 (2010).
2. K. F. Mak and J. Shan, Nat. Photonics **10** 216 (2016).
3. D. Jariwala, V. K. Sangwan, L. J. Lauhon, T. J. Marks, and M. C. Hersam, ACS Nano **8** 1102 (2014).
4. B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, Nat. Nanotechnol. **6** 147 (2011).
5. Y Yoon, K Ganapathi, S Salahuddin, Nanolett **11** 3786 (2011)
6. D. Sarkar, X. Xie, W. Liu, W. Cao, J. Kang, Y. Gong, S. Kraemer, P. M. Ajayan, and K. Banerjee, Nature **526** 91 (2015).
7. C. Gong, H. J. Zhang, W. H. Wang, L. Colombo, R. M. Wallace, K. J. Cho, Appl Phys Lett **103** 053513 (2013).
8. W. Zhou, X. Zou, S. Najmaei, Z. Liu, Y. Shi, J. Kong, J. Lou, P. M. Ajayan, B. I. Yakobson, and J. C. Idrobo, Nano Lett. **13** 2615 (2013).
9. J. Hong, Z. Hu, M. Probert, K. Li, D. Lv, X. Yang, L. Gu, N. Mao, Q. Feng, L. Xie, J. Zhang, D. Wu, Z. Zhang, C. Jin, W. Ji, X. Zhang, J. Yuan, and Z. Zhang, Nat. Commun. **6** 6293 (2015).
10. X Zou, Y Liu, B I Yakobson, Nanolett **13** 253 (2013).
11. S Najmaei, Z Liu, W Zhou, X L Zou, G Shi, S D Lei, B I Yakobson, P M Ajayan, J Lou, Nat Mater **12** 754 (2013).
12. M. Amani, D. Lien, D. Kiriya, J. Xiao, A. Azcatl, J. Noh, S. R. Madhvapathy, R. Addou, S. Kc, M. Dubey, K. Cho, R. M. Wallace, S. Lee, J. He, J. W. A. Iii, X. Zhang, E. Yablonovitch, and A. Javey, Science **350** 1065 (2015).
13. K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, Phys. Rev. B **85**, 115317 (2012).
14. A Allain, J Kang, K Banerjee, A Kis, Nat Mat **14** 1195 (2015).
15. J. H. Park, A. Sanne, Y. Guo, M. Amani, K. Zhang, H. C. P. Movva, J. A. Robinson, A. Javey, J. Robertson, S. K. Banerjee, and A. C. Kummel, Sci. Adv. **3** e1701661 (2017).
16. Z. Yu, Y. Pan, Y. Shen, Z. Wang, Z.-Y. Ong, T. Xu, R. Xin, L. Pan, B. Wang, L. Sun, J. Wang, G. Zhang, Y. W. Zhang, Y. Shi, and X. Wang, Nat. Commun. **5** 5290 (2014).
17. M. Makarova, Y. Okawa, and M. Aono, J. Phys. Chem. C **116** 22411 (2012).
18. K. Cho, M. Min, T. Y. Kim, H. Jeong, J. Pak, J. K. Kim, J. Jang, S. J. Yun, Y. H. Lee, W. K. Hong, and T. Lee, ACS Nano **9** 8044 (2015).
19. D. M. Sim, M. Kim, S. Yim, M. J. Choi, J. Choi, S. Yoo, and Y. S. Jung, ACS Nano **9** 12115 (2015).
20. H Nan, Z Wang, W Wang, Z LIang, Y Lu, Z H Ni, ACS Nano **8** 5738 (2014).
21. P. K. Gogoi, Z. Hu, Q. Wang, A. Carvalho, D. Schmidt, X. Yin, Y. H. Chang, L. J. Li, C. H. Sow, A. H. C. Neto, M. B. H. Breese, A. Rusydi, and A. T. S. Wee, Phys. Rev. Lett. **119** 077402 (2017).
22. Y. Liu, P. Stradins, and S. H. Wei, Angew. Chemie Int. Ed. **55** 965 (2016).
23. M. Amani, R. A. Burke, X. Ji, P. Zhao, D. H. Lien, P. Taheri, G. H. Ahn, D. Kirya, J. W. Ager, E. Yablonovitch, J. Kong, M. Dubey, and A. Javey, ACS Nano **10** 6535 (2016).
24. H. Kim, D. H. Lien, M. Amani, J. W. Ager, and A. Javey, ACS Nano **11** 5179 (2017).
25. A. Alharbi, P. Zahl, and D. Shahrjerdi, Appl. Phys. Lett. **110** 33503 (2017).
26. K L Brower, S Myers, Appl. Phys. Lett. **57** 162 (1990).
27. A H Edwards, Phys. Rev. B **44** 1832 (1991).

28. J H Stathis, E Cartier, Appl. Phys. Lett. **63** 1510 (1993)
29. C J Sandroff, R N Notenburg, J C Bischoff, R Bhat, Appl. Phys. Lett. **51** 33 (1987).
30. M Radosavjevic, et al, Tech Digest IEM (IEEE) (2009)p13.1.
31. L Lin, J Robertson, Appl. Phys. Lett. **98** 082903 (2011).
32. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, Zeitschrift Für Krist. **220** 567 (2005)
33. S. J. Clark and J. Robertson, Phys. Rev. B **82** 085208 (2010).
34. S. Grimme, J Comput Chem **27** 1787 (2006)
35. S. Lany and A. Zunger, Phys. Rev. B **78** 235104 (2008).
36. P. A. Young, J. Phys. D. Appl. Phys. **1** 936 (1968).
37. D. Liu, Y. Guo, L. Fang, and J. Robertson, Appl. Phys. Lett. **103** 183113 (2013).
38. H Lu, Y Guo, J Robertson, Appl. Phys. Lett. **112** 062105 (2018); Y Guo, J Robertson, Phys Rev Mater. **1** 044004 (2017)
39. S. N. Suarez, J. R. P. Jayakody, S. G. Greenbaum, T. Zawodzinski, and J. J. Fontanella, J. Phys. Chem. B **114** 8941 (2010).
40. J. Y. Noh, H. Kim, and Y. S. Kim, Phys. Rev. B. **89** 205417 (2014).
41. C. Tsai, F. Abild-Pedersen, and J. K. Nørskov, Nano Lett. **14** 1381 (2014).
42. H. Li, C. Tsai, A. L. Koh, L. Cai, A. W. Contryman, A. H. Fragapane, J. Zhao, H. S. Han, H. C. Manoharan, F. Abild-Pedersen, J. K. Nørskov, and X. Zheng, Nat. Mater. **15** 48 (2016).
43. M. Amani, P. Taheri, R. Addou, G. H. Ahn, D. Kiriya, D. H. Lien, J. W. Ager, R. M. Wallace, and A. Javey, Nano Lett. **16** 2786 (2016).
44. K. Xiong, J. Robertson, and S. J. Clark, J. Appl. Phys. **99** 44105 (2006).
45. D. Liu and J. Robertson, Appl. Phys. Lett. **94** 42904 (2009).
46. D. Schomburg and G. Michal, John Wiley & Sons, 2012.
47. D. Schomburg, M. Salzmann and D. Stephan, Springer Berlin Heidelberg: Berlin, Heidelberg, 1994; pp 249–251.
48. J. Zhao, K. W. M. Siu, A. C. Hopkinson, Phys. Chem. Chem. Phys. **10** 281 (2008).
49. T. Moehl, M. A. El. Halim, H. Tributsch, J. Appl. Electrochem. **36** 1341 (2006).
50. A. Kutt, T. Rodima, J. Saame, E. Raamat, V. Maemets, I. Kaljurand, I. A. Koppel, R. Y. Garlyauskayte, Y. L. Yagupolskii, L. M. Yagupolskii, E. Bernhardt, H. Willner and I. Leito, J. Org. Chem. **76** 391 (2011).

Chapter 8 Conclusion and Perspectives

This chapter serves as the summary of this thesis and the outline of future work.

8.1 Conclusion

In this thesis, we use DFT to calculation electronic properties of 2D materials as building blocks of the future electronic device. All of the research topics in this thesis are served to one purpose, to enhance the 2D material based device performance, mainly on mobility and conductivity.

The basic knowledge of each property of the different 2D material is illustrated in chapter 1. While in chapter 2 the DFT background theory is stated in detail. In chapter 3 and 4 we calculated one of the most important properties in device physics, the metal-semiconductor contact, and their Schottky barrier height. For defect-free hexagonal nitrides, a chemical trend is shown that Fermi level pinning is relevant to the chemical environment in the interface and the bending of h-XN, both of which are related to the in-plane stability of the 2D materials. The pinning factor can be tuned via inserting oxide layers so the Schottky barrier height can be controlled to a favorable value.

To enhance device performance, besides lowering the contact resistance we need to increase the conductivity of electrodes and the semiconductor channel. Chapter 5 provides a way to increase the carriers' density without degrading their mobility, charge transfer doping. Although graphene has high mobility, the density of carrier is low. It is found that using layers of extreme work function materials can effectively shift the Fermi level off Dirac point. The inter-layer force between graphene and most of the dopants is van der Waals force. Therefore, there is no bending or buckling of graphene and carriers has the same high mobility of as in suspended graphene.

In chapter 6 we pay the attention to defects, which are the main cause of the discrepancy of predicted high mobility and low mobility in reality of a material. We investigate two 2D materials HfS_2 and SnS_2 which are potential building blocks of TFET, which is a low energy consumption transistor requiring an instant response so high mobility is essential. We found that S vacancy and S interstitial and metal interstitial are most significant defects. Besides, these two materials have good band alignment to WSe_2 which is p-type part of TFET. Although HfS_2 and SnS_2 have indirect gaps in 2D, their mobilities are better than MoS_2 and they can compete with InSe as n-type part of TFET.

Beyond studying the type of defects, we can also study how to make the defect chemically inactive, called defect passivation. Chapter 7 illustrate the exact mechanism of how the most prevalent type of intrinsic defect in mechanic exfoliated 2D MoS_2, S vacancy, is passivated by several of methods. The most effective way is superacid passivation. However, the passivation agent and how it passivates are unclear before this work. We found that three hydrogen symmetrically bonded to dangling Mo in S vacancy, with -1 charge is the only solution. Hydrogen passivates S vacancy by forming bonding and anti-bonding state with defect states so all gap states are cleaned and Fermi level is in the middle of the gap.

8.2 Future work

1. The metal-insulator transition (MIT) of VO_2 will be investigated. VO_2 is in rutile non-magnetic form in high temperature while when the temperature reduces below 340K it changes to M1 phase which is anti-ferromagnetic and monoclinic form. It can be used as smart windows [1, 2]. In this work, we will study how doping affects the transition temperature and magnetic order, as well as the band gap. The GGA+U and HSE functional are used to find the correct geometry and magnetic order. We just found that doping will not break the anti-ferromagnetic order of VO_2 M1 phase. By adding MgO to VO_2 with ratio of 1:2, the gap increases and by adding Ge, the phase transition temperature changes.

2. Carbon nitride (CN) is a promising photo-catalyst in hydrogen evolution reaction (HER) [3-6]. Various carbon nitrides have been successfully synthesized with high hydrogen production rates. However, there is no solid study of CN in the aspect of electronic structure. The aim of this project is to present the density functional theory (DFT) study of graphitic C_3N_4 and defect-rich amorphous carbon nitride (a-CN_x).The amorphous CN structure is created by molecular dynamics with a surface. We found that a-CN_x has smaller band gap of 0.3eV than that of g-C_3N_4, in agreement with available experimental data. The inversion participation ratio is calculated to see the degree of localization of the states near Fermi level. We found a peak near the valence band maximum (VBM). The band diagram calculated yields high-efficiency solar energy conversion.

3. The sulphur vacancy is the primary defect in mechanical exfoliated MoS_2 but it is not the case in chemical vapour deposition (CVD) grown MoS_2 [7, 8]. The Grain boundary is the most important type of the defect and it brings about defect states in the gap as well. We will use halide atoms, small organic molecules and hydrogen to passivate the grain boundary. Apart from MoS_2, we also found H can passivate 2D SnS_2.

REFERENCES

1. S Wang, M Liu, L Kong, Y Long, X Jiang and A Yu, Prog. Mat. Sci. **81**, 1 (2016).
2. Y Cui, Y Ke, C Liu, Z Chen, N Wang, L Zhang, Y Zhou, S Wang, Y Gao and Y Long, Joule (2018).
3. K Schwinghammer, M B Mesch, V Duppel, et al. J. Am. Chem. Soc. **136** 1730 (2014).
4. J H Zhang, X F Chen, K Takanabe, K Maeda, K Domen, J D Epping, X Z Fu, M Antonietti, X C Wang, X. C. Angew. Chem., Int. Ed. **122**, 451 (2010).
5. S Yang, Y Gong, J Zhang, L Zhan, L Ma, Z Fang, R Vajtai, X Wang, P M Ajayan, Adv. Mater. **25**, 2452 (2013).
6. X Chen, Y S Jun, K Takanabe, K Maeda, K Domen, X Fu, M Antonietti, X Wang, Chem. Mater. **21**, 4093 (2009).
7. V Zande, M Arend, et al. Nat. Mater. **12** 554 (2013).
8. L Karvonen, A Saynatjoki, M J Huttunen, A Autere, B Amirsolaimani, S Li, R A Norwood, N Peyghambarian, H Lipsanen, G Eda, K Kieu and Z Sun, Nat. Comm. **8** 15714 (2017).